

Toward a Strong-Sense Validation Benchmark Database for Numerical Ladle Flow Models



TIM HAAS, ANA LIA SUAREZ, MORITZ EICKHOFF, and HERBERT PFEIFER

In numerical computation of ladle metallurgy, multiphase models are essential. Still, these models are afflicted with great uncertainty, making a validation with experimental data mandatory. Validation experiments fundamentally differ from physical modeling experiments because emphasis is on a complete documentation of all boundary conditions and a detailed uncertainty assessment. For this work, the experimental design for a comprehensive accuracy assessment of numerical models and reported quantities were jointly determined with international numerical experts. The framework comprises a plume analysis and flow measurements in the single- and multiphase regions of a water model with defined conditions. All influencing factors are documented extensively and their contribution to the overall data uncertainty is quantified. Detailed data are made publicly available within a validation benchmark database for isothermal multiphase flow in metallurgical ladles. A first draft of a standardized validation procedure, including a single number validation score, is proposed. Using the available data, the accuracy of numerical models can be quantified more accurately and comparably, which helps in advancing the model's further development. It also lays the foundation for a standardization of the validation process, which can lead to greater acceptance of the numerical results.

<https://doi.org/10.1007/s11663-020-01987-6>
© The Author(s) 2020

I. INTRODUCTION

LADLE metallurgy is an important process in steel refining. In the ladle, a homogenization of temperature and composition is achieved by gas purging while the introduced bubbles promote the removal of nonmetallic inclusions. However, the tolerable levels of inclusions decrease steadily, which constantly raises the demands on process control. Owing to that, an increasingly detailed knowledge of all process quantities becomes crucial.

Temperatures of about 1600 °C, large scales, and the visual opacity of liquid steel make direct measurements of most process parameters impossible in steel metallurgy. An established alternative to access these quantities is numerical models. However, with increasing process requirements, the demanded levels of accuracy and detail are rising. A particular challenge is the occurrence of different interacting phases, which require the use of multiphase models. Despite numerous

improvements in recent years, these are still associated with a great deal of uncertainty. This makes validation of the numerical model mandatory.

Because of the mentioned difficulties of direct measurements in the actual process, different measurement methods, mainly in water models, were used for this purpose. An overview of different numerical models for the fluid flow in the ladle and their validation methods is given in Table I.

This summary illustrates that most researchers used literature data from three distinctive experiments.

Castillejos and Brimacombe^[13] used a double-contact electroresistivity sensor in slightly acidified, deionized water to measure the local gas fraction, bubble frequency, bubble rising velocity, and diameter of the plume. Apart from ladle metallurgy, their data were also used in some fundamental works^[40–42] on the numerical modeling of gas bubbles.

Sheng and Irons^[4] measured the bubble velocity and the liquid velocity in the plume region by combining laser doppler anemometry (LDA) and magnetic probe in the plume zone of a water model. LDA was used to measure the local velocity components, while magnetic probe was used to determine whether the recorded signal belonged to a bubble or to the bulk liquid. Their experiments yield the gas and liquid velocities, a void fraction distribution, and the turbulent kinetic energy (TKE) derived by the flow measurements.

TIM HAAS, ANA LIA SUAREZ, MORITZ EICKHOFF, and HERBERT PFEIFER are with the Department for Industrial Furnaces and Heat Engineering, RWTH Aachen University, Kopernikusstraße 10, 52074 Aachen, Germany. Contact e-mail: haas@iob.rwth-aachen.de

Manuscript submitted March 19, 2020; accepted September 29, 2020.

Article published online November 3, 2020.

Mietz and Oeters^[43] used LDA in a water model to obtain the general flow field in the ladle for centric and eccentric positioned nozzles.

Xie *et al.* conducted measurements in Wood's metal. They used double-contact electroresistivity probes^[26] to measure the local gas fraction, bubble frequency, size distribution, and bubble rising velocity. In a subsequent study, they employed permanent magnet probes to measure the local flow velocity and derived mean velocity fields, TKE, and turbulent dissipation rates in a model with centric nozzle position.^[44] Later, they repeated the measurements for an eccentric nozzle position.^[45]

These studies have in common that they were intended to gain additional knowledge about some flow phenomena, not primarily for the sake of validation. However, validation experiments are fundamentally different from those physical modeling experiments of their main purpose.^[46] They are intended to compare measurement quantities recorded in a well-defined system to a numerical model of exactly the same system. Thus, emphasis is on a detailed characterization and documentation of the experiments, including fluid properties, geometry, boundary conditions, and measurement uncertainties. A characteristic of validation experiments is that their quality is determined by the level of completeness of provided documentation. To understand the importance of detailed and comprehensive documentation, one can imagine the numerical modeling of an experiment that is described insufficiently. Any deviation between the numerical model and the experiments can be caused either by inaccurate boundary conditions or by inadequate mathematical models. It is almost impossible to assign the deviation to one of these sources. This prevents a structured optimization of the mathematical model. Instead, there is a danger of worsening mathematical submodels to adapt the results to wrong boundary conditions. According to Oberkamp and Trucano,^[47] a thorough documentation should contain a comprehensive conceptual and experimental description, an uncertainty quantification of the measurement results, and all additional information that might be needed by the users of the validation experiment. To accomplish this, the measurements should be designed to analyze experimental uncertainties. Oberkamp and Smith^[48] divided the uncertainties into aleatory and epistemic uncertainties. Aleatory uncertainties are model intrinsic. An example is that the flow can settle at slightly different stable states. Epistemic uncertainty is caused by missing data in the documentation or measurement errors. In addition, it is useful to distinguish between random and systematic uncertainties. In contrast to systematic uncertainties, the effect of random uncertainties decreases with the number of samples. Aleatory uncertainties are always random uncertainties, while epistemic uncertainties can be systematic or random. A distinction between the uncertainties can be difficult to accomplish.

For the structure of a validation benchmark database, Oberkamp and Smith^[48] proposed the idea of a validation hierarchy that ranges from complete system tiers to unit problem tiers with different subtiers. The

different tiers differ in the level of computational complexity and their main purpose. While higher-level tiers can be used to validate the numerical model, low-level tiers can be used for model calibration. The complete system is usually too complex to be validated directly. Instead, the validation hierarchy allows division of the system into different subsystems that can be validated. However, it is worth noting that not all subsystems can be validated independently but that the accuracy of the other subsystems needs to be considered.

Best practice guidelines for the design and execution of validation benchmark experiments were summarized by Oberkamp and Smith.^[48] In particular, they emphasized that the experiments should be designed in close collaboration between the experimentalists and developers of numerical models. By that, the validation experiments should be suitable for a comprehensive accuracy assessment of the numerical model. In addition, this should ensure that the documentation contains all information necessary for the complete representation of the setup in the numerical model, including all initial and boundary conditions and external influences. However, the experiments and numerical calculations should be performed independently of each other to avoid eventual biases.

Compliance with these guidelines and a thorough documentation are prerequisites for strong-sense validation benchmark databases, which is a concept introduced by Oberkamp *et al.*^[49] A strong-sense benchmark is defined by the following:

- (1) an exact, standardized, frozen, and promulgated definition of the benchmark;
- (2) an exact, standardized, frozen, and promulgated statement of the purpose of the benchmark (this statement addresses the benchmark's role and application in a comprehensive test plan for a code, for example);
- (3) exact, standardized, frozen, and promulgated requirements for comparison of codes with the benchmark's results; and
- (4) an exact, standardized, frozen, and promulgated definition of acceptance criteria for comparison of code with the benchmark's results (these criteria can be phrased either in terms of success or failure).

According to Oberkamp *et al.*,^[49] strong-sense benchmarks do not exist hitherto and the establishment of such standards will be a long and difficult process. However, the establishment of such benchmarks and standards would offer great advantages for the optimization and acceptance of numerical models.

In a recent work, Haas *et al.*^[50] compared different validation methods for the isothermal flow in a ladle and critically discussed whether they are suitable for evaluating the accuracy of a numerical model. It was shown that an evaluation of the slag eye is not sensitive enough for validation. In fact, a good qualitative agreement between all tested numerical models and the measured slag eye was found, even though a cross-validation with particle image velocimetry (PIV) and bubble tracking revealed decisive inaccuracies. Furthermore, the evaluation of experimental images is usually made manually,

Table I. Numerical Models for the Fluid Flow in the Ladle and Employed Validation Methods

Authors	Year	Validation Method
Sahai and Guthrie ^[1]	1982	flow field (video recordings of small, rectangular cards) (M)
Mazumdar and Guthrie ^[2]	1985	flow visualization (suspended grid work of silken threads) (M)
Joo and Guthrie ^[3]	1992	mixing time (M)
Sheng and Irons ^[4]	1993	averaged velocity field, TKE, void fraction, and bubble frequency by LDA with electric probe; bubble size distribution by digital camera (M)
Mazumdar <i>et al.</i> ^[5]	1993	averaged velocity from Ref. [6]; average bubble rising velocity and vertical velocity of the plume from Ref. [7]; axial velocity from Ref. [1] (M)
Goldschmit and Owen ^[8]	2001	axial liquid velocity from Ref. [4] (M)
Ganguly and Chakraborty ^[9]	2004	temperature distribution and averaged velocity field from Ref. [10] (FS)
Aoki <i>et al.</i> ^[11]	2004	mixing time (FS)
Mendez <i>et al.</i> ^[12]	2005	void fraction from Ref. [13] (M)
Singh and Mazumdar ^[14]	2008	axial liquid velocity from Ref. [4]; shear stress from Ref. [15] (M)
Cloete <i>et al.</i> ^[16]	2009	axial liquid velocity and TKE from Ref. [4] (M)
Llanos <i>et al.</i> ^[17]	2010	slag eye and mixing time (M)
Geng <i>et al.</i> ^[18]	2010	mixing time from Ref. [3] (M)
Maldonado-Parra <i>et al.</i> ^[19]	2011	temperature distribution from Ref. [20] (FS)
Liu <i>et al.</i> ^[21]	2011	slag eye opening area from Refs. [22] through [24] (M)
De Felice <i>et al.</i> ^[25]	2012	void fraction and vertical liquid velocity from Ref. [26] (M)
Lou and Zhu ^[27]	2013	void fraction from Ref. [13]; axial liquid velocity and TKE from Ref. [4] (M)
Huang <i>et al.</i> ^[28]	2013	—
Yu and Louhenkilpi ^[29]	2013	axial liquid velocity and TKE ^[4] (M)
Bellot <i>et al.</i> ^[30]	2014	mixing time; deoxidation rate from Ref. [31] (M)
Li <i>et al.</i> ^[32]	2015	slag eye opening area and bubble Sauter mean diameter (M)
Li <i>et al.</i> ^[33]	2017	slag eye opening area (M)
Haas <i>et al.</i> ^[34]	2017	flow field by PIV (M)
Duan <i>et al.</i> ^[35]	2018	axial liquid velocity and bubble volume fraction from Ref. [4] (M)
Ramasetti <i>et al.</i> ^[36]	2018	slag eye opening area (FS)
Ramasetti <i>et al.</i> ^[37]	2019	slag eye opening area from Refs. [22] through [24] (M); slag eye opening area (FS); mixing time (FS)
Alia <i>et al.</i> ^[38]	2019	axial liquid velocity from Ref. [39] (M)

TKE turbulent kinetic energy, *LDA* laser doppler anemometry, *PIV* particle image velocimetry, *FS* full scale, *M* physical model.

which allows neither reproducibility nor comparability. Bubble swarm tracking is a suitable method to assess the accuracy of bubble-related submodels, while it does not fully address the impact on the main flow field. In contrast to that, flow measurements of the bulk provide an appropriate evaluation of the overall flow structure but can be misleading if shown as line plots. In addition, they are not sensitive enough to assess the accuracy of all submodels. As a general rule, it is not recommended to use only one type of measurement to avoid overfitting the validation data. Based on these results, a two-step best-practice validation strategy can be derived. In the first step, a validated optimization of the numerical model for the isothermal flow in a water model should be carried out. Flow measurements using PIV and an analysis of the plume region should be combined to assess different submodels simultaneously. However, since the water model does not consider thermal effects, the influence of slag, or inhomogeneous chemical composition, the model's upscaling to industrial conditions should be cross-validated with measurements in the actual ladle in the second step. For this purpose, either measurements of the slag eye,^[37] vibration of the vessel,^[51] or the concentration of alloying elements^[11,52] might be used. However, a critical analysis of the methods regarding their suitability for validation is still pending. Splitting the validation into a two-step procedure is necessary because plant data do not provide

sufficient detail to perform model optimization and the dependence on certain submodels leads to a high tendency of overfitting.

To lay the foundation of a standardized validation procedure that follows that two-step strategy, validation data for the first step are provided in this work. Validation experiments are conducted to assess the accuracy of numerical models for the isothermal flow in the ladle. In order to approach the long-term goal of a strong-sense benchmark for steel metallurgy, the results are collected in a publicly accessible validation database.^[53] Here, the concept of the database is described and comprehensive documentation of the experimental setup, measurement procedures, and further influencing factors is provided. This includes a detailed analysis of the measurement uncertainty for the provided data. Finally, guidelines for the use of the provided data are defined and the concept of a validation score is introduced.

A. Database Concept and Structure

Following the guidelines for the design and execution of validation benchmark experiments by Oberkamp and Smith,^[48] the concept of the database was jointly developed with international computational fluid dynamics (CFD) experts in the field of metallurgy, who are listed in the Acknowledgments. The scope was

to identify the necessary data for a sensitive and comprehensive though detailed accuracy assessment. A requirement catalog regarding the documentation of the measurements was defined.

In addition to bubble swarm tracking and the mean flow field derived by PIV, an analysis of the fluctuation velocity, the flow in the plume region, and closeups of areas of particular importance were suggested. A detailed analysis of the bubble size distribution was included either as a model input or as a validation for population balance models (PBMs). For all measurements, different sources of uncertainty were discussed and the experiments were designed to quantify most of them.

The general structure of the experimental design is given in Figure 1. It is structured as a validation hierarchy following the example of Oberkampff and Smith.^[48] That system allows the addition of other systems to the subsystem tier, such as plant data, in the future.

A description of all experiments, including a definition of the setup and all boundary conditions as well as an analysis of the accuracy, is given in this work. Additional information that is too detailed to be described here is provided for each branch. All presented results are derived from experiments with a gas flow rate of 2.4 slm. In addition, the database also contains the same measurements for flow rates of 1.2, 1.8, and 3 slm.

The documentation of the experiments follows the guidelines defined together with the experts. For all data that could not be quantified, a qualitative estimate was made based on justified assumptions. The level of detail of the documentation exceeds that of physical modeling experiments by far. Some details cannot yet be considered in a CFD model or have no influence on the results according to the current state of knowledge. However, it can be assumed that numerical models will become more detailed in the future and that not all influencing parameters are fully understood yet.

II. EXPERIMENTAL SETUP

All measurements are carried out in a slightly simplified acrylic glass (polymethyl methacrylate) model of an industrial 185 t ladle, geometrically scaled by 1:5 (scale factor $\lambda = 0.2$). While the walls are straight, the bottom is slightly rounded. Due to production, the bottom is not completely smooth but contains steps (maximum 5 mm) at the places where porous plug plates can be installed. The ladle model is placed in an outer water tank that minimizes optical distortions and stabilizes the walls mechanically. The dimensions are marked in Figure 2. Additional technical drawings of all parts and images of the model and the setup can be found in the database.

A dynamic similarity criterion for the isothermal, homogeneous, and slag-free flow can be derived from the momentum equation. Considering the ratios of its relevant terms, the inertial force, the viscous force, the gravitational force, and the pressure force, dynamic similarities between the flow in the real vessel and the model are maintained by a similarity of the nondimensional Reynolds and Froude numbers. However, in

geometrically downscaled models, it is impossible to meet both requirements. The Reynolds number is usually very high in the range of 10^5 . Consequently, it can be concluded that the influence of the viscous force is very small compared to the inertial force. Thus, dynamic similarity is determined by a Froude number criterion. For the validation experiments, the plume Froude number criterion by Krishnapisharody and Irons^[54] is used so that the flow rate is scaled by

$$\dot{Q}_M = \lambda^{2.5} \dot{Q}_{FS} \quad [1]$$

where \dot{Q}_M is the flow rate in the model, \dot{Q}_{FS} the flow rate in the industrial ladle, and λ the scaling factor. The investigated flow rates and their correspondence in the actual process are given in Table II.

With the given geometric factor, the Reynolds number in the model is about one order of magnitude smaller than in the real vessel. Thus, the effect of viscous forces is overestimated in the downscaled water model. However, it is still sufficiently high so that the effect of viscous forces can be considered negligible compared to the inertial force.

The model is about 10 years old and was frequently used. It was not exposed to direct sunlight. Thus, it can be concluded that the acrylic glass contains microcracks and some larger scratches from use. Unfortunately, there are no techniques available in the lab to quantify the effects on the actual surface roughness, wettability, or mechanical, optical, and thermal properties of the material.

Gas is injected through a porous plug. More specific details about the hoses and connections are provided in the database. The gas flow rate is controlled by a digital mass flow controller (ANALYT-MTC 35833, ± 0.045). A gage pressure of 2 bar is maintained (Riegler 0.2 to 6 bar) at the inlet of the mass flow controller. Pressurized air is provided by a compressor with integrated dryer (Atlas Copco GX4 FF). The gas temperature above the porous plug without water is 19 °C, measured with a digital thermometer. The porous plug is a radial distance of 0.21 m from the ladle axis and has a diameter of 0.02 m. The porous plug's original properties are summarized in Table III. Note that the frequent contact with water and tracer particles might have changed the effective properties. Unfortunately, that effect cannot be quantified.

To reduce thermal gradients, the water temperature is kept at $20 \text{ °C} \pm 0.2 \text{ °C}$, which approximately matches the ambient air temperature. The temperature is adjusted by mixing regular tap water (14.5 °C) and warm tap water (41 °C) during the filling. After filling, gas purging is started and is maintained for 10 minutes to assure that the flow has come to a settled state and the temperature is homogenized. Throughout the experiments, the temperature decreases at a rate of about 0.1 °C/h. Thus, the assumption of an isothermal flow is justified. To account for the energy loss, the temperature is controlled by mixing warm water. The temperature is homogenized by gas stirring for 5 minutes before the next measurement is conducted. During the filling process, a hole in the bottom of the model is opened

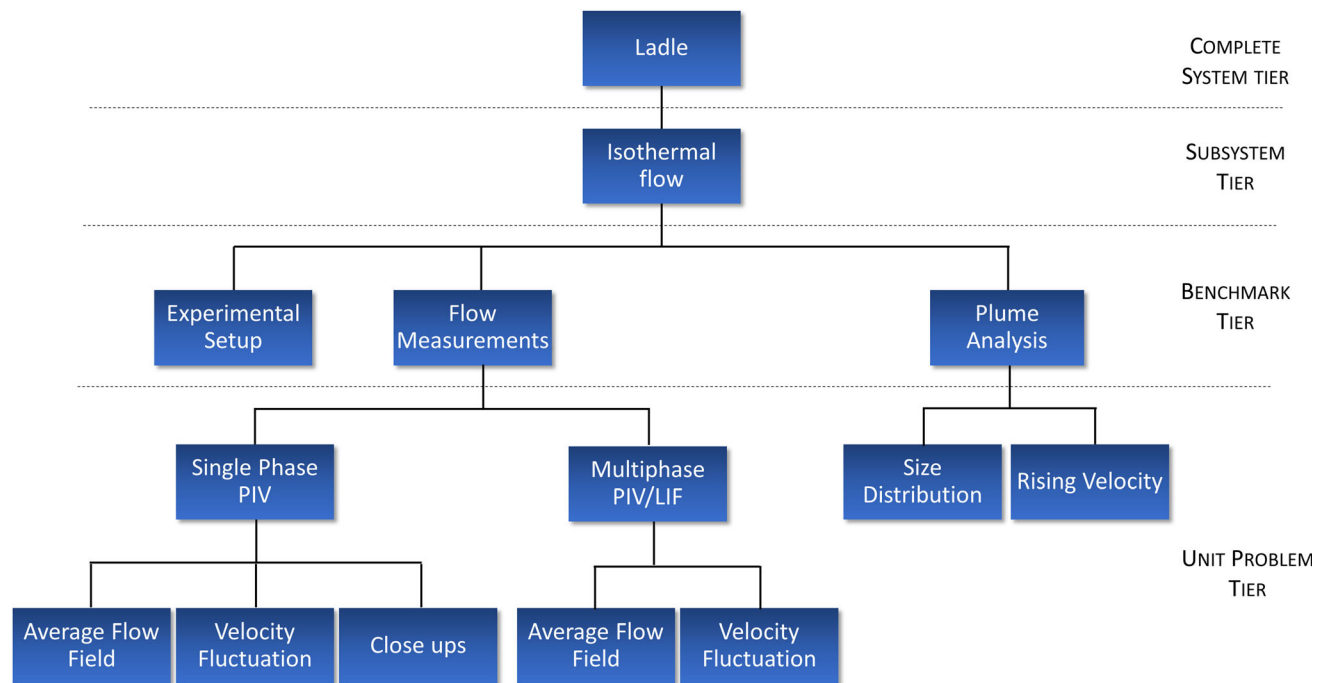


Fig. 1—Validation hierarchy and structure of the experimental design for the validation database.

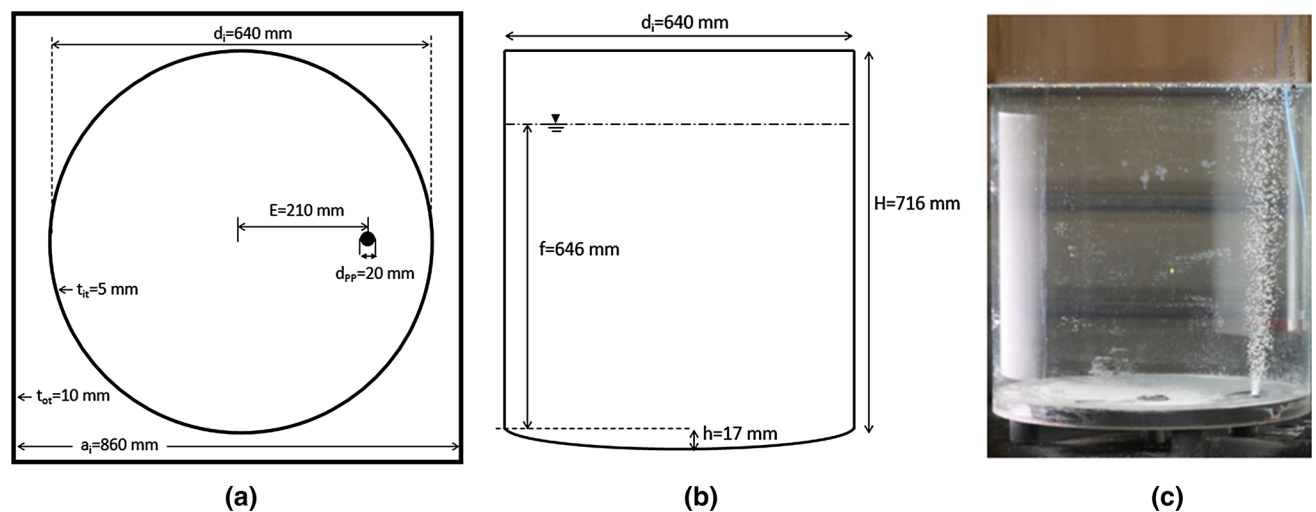


Fig. 2—Acrylic glass model of a 185 t ladle: drawing of (a) top view and (b) side view and (c) real setup.

Table II. Gas Flow Rates in the Validation Database and Corresponding Plant Flow Rates

	Gas Flow Rate (L/min)			
	1.2	1.8	2.4	3.0
Water Model	67.1	100.6	134.2	167.7
Real Ladle				

to ensure that the filling height in the inner and outer tanks is exactly the same. Additional details about the water hardness and composition can be found in the database.

A. Uncertainty Quantification

Oberkamp and Smith^[48] pointed out that the level of completeness of provided documentation can be a major source of uncertainty for validation experiments because it causes a deviation between the numerical system and the experiment. Documentation incompleteness can have two major sources. First, a boundary condition is known to be important but not all details can be provided. That is usually the case when specific measurement techniques are unavailable or if no measurement procedure has been established yet. This applies to the actual conditions of the porous plug or the surface quality of the acrylic glass. Only information about their

initial state is available, but it is likely that this state has changed over its lifetime. The amount of dissolved gases or contaminants in the water are also known to have an effect, but they cannot be quantified. It is the authors' opinion that this information is not crucial for the evaluation of current numerical models. However, in the future, this might become a serious drawback for more detailed models. The second source of incompleteness is that influencing factors are missed or not identified as such yet. Identifying all relevant boundary conditions is a major challenge in the conduction of validation experiments. The validation database was collaborated with international CFD experts to minimize the effect of missing boundary conditions.

B. Plume Analysis

Different techniques are in use to quantify plume characteristics, such as the bubble rising velocity and the bubble diameter. For the purpose of validation, electroresistivity probes have mainly been used, though in more recent studies, imaging has been employed as well. In this work, imaging in combination with digital image processing is applied. This technique may make evaluation more difficult than electroresistivity probes, because it might introduce some uncertainties during evaluation. On the other hand, it is nonintrusive, that is to say it reduces the systematic uncertainty because the sensor does not interact with the measured quantity.

C. Experimental Procedure

The rising velocity of ascending bubbles in the bubble swarm and the bubble size distribution are measured by a high-speed camera (Photron FASTCAM SA3, resolution 1024×1024 pixels) using a 60-mm lens and a frame rate of 500 frames per seconds. The shutter speed is set to $1/3000$ seconds. Further settings can be found in the database. Because the image acquisition is faster than the storage on a hard drive, the images are first stored in the internal storage of the camera. This limits the maximum number of images to 5457. To simplify the image analysis, an LED panel as a homogeneous, diffuse backlight is placed behind the bubble plume, as shown in Figure 3.

High-speed images are taken for four different flow rates (1.2, 1.8, 2.4, and 3 slm) at five different heights ($z = 0.12, 0.24, 0.36, 0.48,$ and 0.60 m). For each height, the camera is calibrated before image acquisition for the different flow rates. After the flow rate is adjusted, it takes 5 minutes to ensure that the flow has reached a stable state. The water is seeded with PIV tracer particles (Vestosint 1111, $\rho = 1.016$ g/cm³,

$d_{\min} = 50 \mu\text{m}$, $d_{\max} = 75 \mu\text{m}$) because it is known that particles can have an impact on the rising velocity, bubble shape, and coagulation and break-up behavior.^[55] By keeping the tracer concentration constant throughout all experiments, including the PIV measurements, it is ensured that the results can be compared with the same numerical model.

The conversion from pixels to a metric scale is conducted with an automated procedure. An image of an equidistant checkboard is taken and the touching points of the squares are automatically detected based on local intensity gradients. Afterward, the distance to the four closest points is evaluated for each point. Distances further than 1.2 times the median distance are assumed to belong to corner points that have less than four checkerboard neighbors. Finally, the mean calibration factor is derived by dividing the mean distance in pixels by the known square width of the actual calibration plate.

D. Data Processing

The bubble's rising velocity and diameter can be derived from the videos through different techniques, which are discussed subsequently. By employing different techniques, the uncertainty introduced by the image processing can be estimated.

1. Manual evaluation

Manual evaluation of the bubble rising velocity is by far the simplest method. Here, all bubbles are marked manually with the cursor. Either an ellipse is created and manually adjusted until it describes the bubble shape or the major and minor axes of the bubble are marked. Afterward, the center of the bubble has to be marked in a few consecutive frames. The rising velocity is computed by the displacement of the bubble over the frames. By that, information about the bubble size and the rising velocity are derived concurrently. A major advantage of the method is its simplicity. There are almost no requirements or restrictions for equipment and experimental conditions. Decent results can even be obtained by using the high-speed video recording mode of modern smartphones. In addition, the researcher has full control over the results and benefits from human intuition. On the other hand, it is very time consuming and hardly feasible to mark all bubbles in consecutive frames. Because of that, the method usually does not capture time statistics about the plume. Furthermore, the results are not fully reproducible because they depend on the person making the evaluation. Nonetheless, statistically significant mean bubble velocity and size distributions can be derived if enough bubbles are included.

2. Automated bubble detection

An automated analysis of the bubble swarm videos has to be separated in two different tasks. The first is the detection of bubbles on the frames. Different methods for that are available and are discussed briefly subsequently. The second task is to assign the detections found on different frames to a coherent track.

Table III. Porous Plug Properties

Diameter (mm)	20
Open Porosity (Vol Pct)	27
Bulk Density (g/cm ³)	2.6
Raw Material Basis	MgO

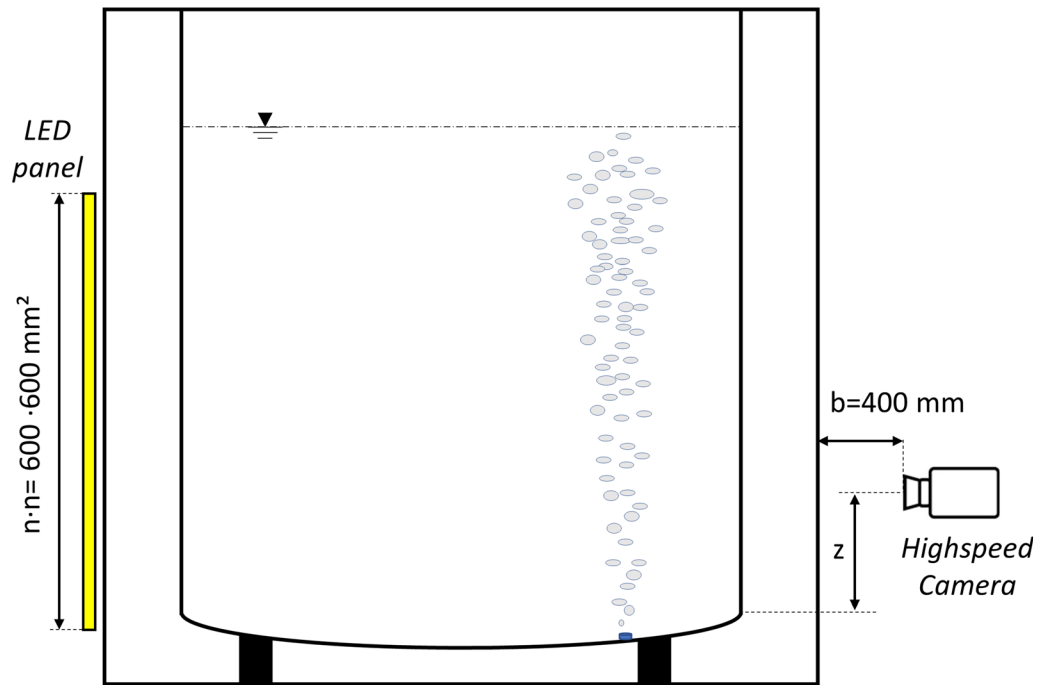


Fig. 3—Experimental setup of bubble swarm tracking.

The bubble swarm in water models of the ladle has an intermediate to high void fraction. Because of that, most bubbles are visible as clusters in the two-dimensional image. A particular challenge in bubble detection is the segmentation of clusters into individual bubbles. The bubble shapes are usually approximated by ellipsoids, which become ellipses in a two-dimensional projection. The assumption of ellipsoid bubbles is valid up to an Eotvos number of about 40, while the deviation increases with increasing Eotvos and Reynolds numbers.^[56] In water models of ladles the bubbles, the Reynolds number is about 750, while the Eotvos number it is about 2. Thus, the assumption is valid for most bubbles in the plume. In contrast to manual clicking, automatic detection is only feasible in the case where images have sufficient quality. Thus, a high-speed camera with a shutter speed of at least 1/1000 seconds and a strong, homogenous backlight is mandatory. Automated detection can either be achieved by conventional image processing or by convolutional neural networks (CNNs).

3. Conventional image processing

Automation can be achieved by employing a multi-stage digital image processing procedure for the detection and reconstruction of single and overlapping bubbles. Generally, these procedures include an object detection stage and an object segmentation stage, where objects that are bubble clusters are segmented into parts of individual bubbles. In the final stage, the bubble shapes are approximated by an ellipse. Different workflows are proposed in the literature.^[57–61] However, a particular drawback is that the workflow and its parameters strongly depend on the experimental setups

in which the images are acquired. Therefore, no generally applicable image processing procedure for bubble detection has been proposed yet. In this work, a method is employed that follows a four-step framework that consists of boundary extraction, concave point detection, boundary segmentation, segment grouping, and contour estimation, as shown in Figures 4 and 5.

First, outer and inner boundaries of single and clustered bubbles are extracted using global and adaptive thresholding. Second, polygonal approximation^[62] is performed on the outer boundaries to find the concave points while assuming that these represent the connecting points of overlapping bubbles. Third, depending on the number of breakpoints found on the object boundary, it is decided if the object corresponds to a single bubble or to a cluster of overlapping bubbles. In the latter case, the outer boundary is split into segments by the concave points. Segments that belong to one bubble are joined together based on the idea that they should be near the same inner bubble. Finally, an ellipse-fitting algorithm using the least-squares criterion is used to approximate the contour of single and clustered bubbles.

4. BubCNN

An alternative approach based on machine learning has been proposed by Haas *et al.*^[63] In contrast to conventional image processing, features are not extracted manually, but the program learns to identify bubbles based on a labeled training data set. The detector, called BubCNN, shown in Figure 6, employs two pretrained modules, a Faster region-based convolutional neural network (RCNN)^[64] and a shape regression CNN. The Faster RCNN module detects an

undetermined number of bubbles on arbitrarily sized images and marks them by bounding boxes (a). After that, image patches of the located bubbles are extracted, resized, and processed by the shape regression CNN that approximates the bubble shape by an ellipse (b).

The program, including two pretrained modules, is publicly available on GitHub: <https://github.com/Tim-Haas/BubCNN>. An additional transfer learning module allows customization of BubCNN to specific experimental conditions in case the pretrained modules do not yield satisfying results.

A major advantage of the machine learning approach is that it is better at generalizing than conventional image processing and that it can be customized without expert knowledge about image processing or machine learning. On the other hand, training is a statistical process, so the results are not fully reproducible and the accuracy varies in a small margin. For the recorded videos in this work, BubCNN v.1.01 is employed. To achieve higher accuracies, the transfer learning module is used to customize the networks to the described experimental setup. For that, all bubbles are marked semiautomatic in two random frames. In addition, assembling is used to decrease the detection recall. By that, two Faster RCNN modules are employed that are previously trained on similar data sets. Because learning is a statistical process, modules yield slightly different detections. Merging these detections decreases the missing rate of the detectors.

Table IV summarizes the advantages and drawbacks of the different techniques. For this work, BubCNN is chosen because it runs faster than conventional image processing on a GPU and can be used for extensive statistical analysis. The other techniques are used to determine the systematic uncertainty introduced by the detection algorithm.

5. Track assignment

In case the bubbles are detected automatically, the detection on the consecutive frames has to be assigned to coherent tracks, which is done by a cost function:

$$J = |\vec{p}_{\text{track}} - \vec{p}_{\text{detection}}| + |a_{\text{track}} - a_{\text{detection}}| + |b_{\text{track}} - b_{\text{detection}}|, \quad [2]$$

where p_{track} is the predicted center of a bubble in a track estimated by a Kalman filter, $p_{\text{detection}}$ is the center of a detection, a_{track} and b_{track} are the axes of the tracked bubble averaged over all previously assigned detections, and $a_{\text{detection}}$ and $b_{\text{detection}}$ are the axes of the detection.

The cost is computed between all active tracks and all detections on the next frame. A detection is assigned to the track with the lowest cost. In cases where all costs exceed 10, the detection creates a new track. Tracks are only considered valid in cases in which a detection is assigned to them at least 10 times and the detected z -position increases over the frames. Otherwise, the track is discarded.

Because the tracks are only two dimensional, the rising velocity is defined as the z -velocity of the bubbles. It is computed by

$$\overline{w_b} = \frac{\overline{\Delta z} \cdot M}{\Delta t}, \quad [3]$$

where $\overline{w_b}$ is the averaged bubble rising velocity, $\overline{\Delta z}$ is the mean z distance between all consecutive detections assigned to a track, and M is a scaling factor. Since the position of the calibration plate can slightly change at different calibrations, the plume center is defined as the mean x position averaged over all tracks.

To obtain plume rising profiles, the tracks are assorted into gridded categories based on their position relative to the center of the plume and their averaged rising velocity. The position categories range from -67.5 to 67.5 mm in steps of 5 mm. The velocity categories range from 0 to 1 m/s in steps of 0.05 m/s. The average rising velocity for each position category is computed by summing the products of the number of bubbles in a velocity category by its velocity value and dividing this sum by the number of tracks assigned to all velocity categories for the specified position category.

It should be noted that the rising profiles, especially the values further off the plume center, depend to some extent on the range and spacing of the categories. Therefore, these categories have to be chosen to meet the requirements of the plume. However, the plume characteristics of the numerical model are not known *a priori*. Therefore, the mean bubble rising velocity and the width of the plume are used as validation criteria. For that, the plume width is defined as the length that contains 50 pct of the bubbles closest to the plume center. These criteria are less sensitive than the bubble rising profiles; on the other hand, they are also less sensitive to postprocessing. In the future, more accurate numerical models might require the profiles as validation criteria.

For the validation of PBMs, the equivalent diameter is derived. For that, the mean of both semiaxes of all detections assigned to a track are computed. Based on the assumption that the ellipses are rotationally symmetric along the minor axis, the equivalent diameter can be computed by

$$d_{\text{Eq}} = 2 \cdot (d_{\text{min}} d_{\text{max}}^2)^{1/3}, \quad [4]$$

where d_{Eq} is the equivalent diameter, d_{min} is the minor ellipse axis, and d_{max} is the major axis. The factor 2 has to be multiplied because BubCNN yield in semi-axes is defined as the distance between the bubble center and its outline.

E. Uncertainty Quantification

It is examined whether the high-speed recordings are sufficiently long so that the epistemic uncertainty caused by the number of samples becomes small, and it is found that the last 50 tracks change the mean bubble rising velocity and the plume width by less than 0.005 pct. To test the aleatory uncertainty, three experiments with the same flow rate, captured at the same height, are conducted. The mean relative deviation between the

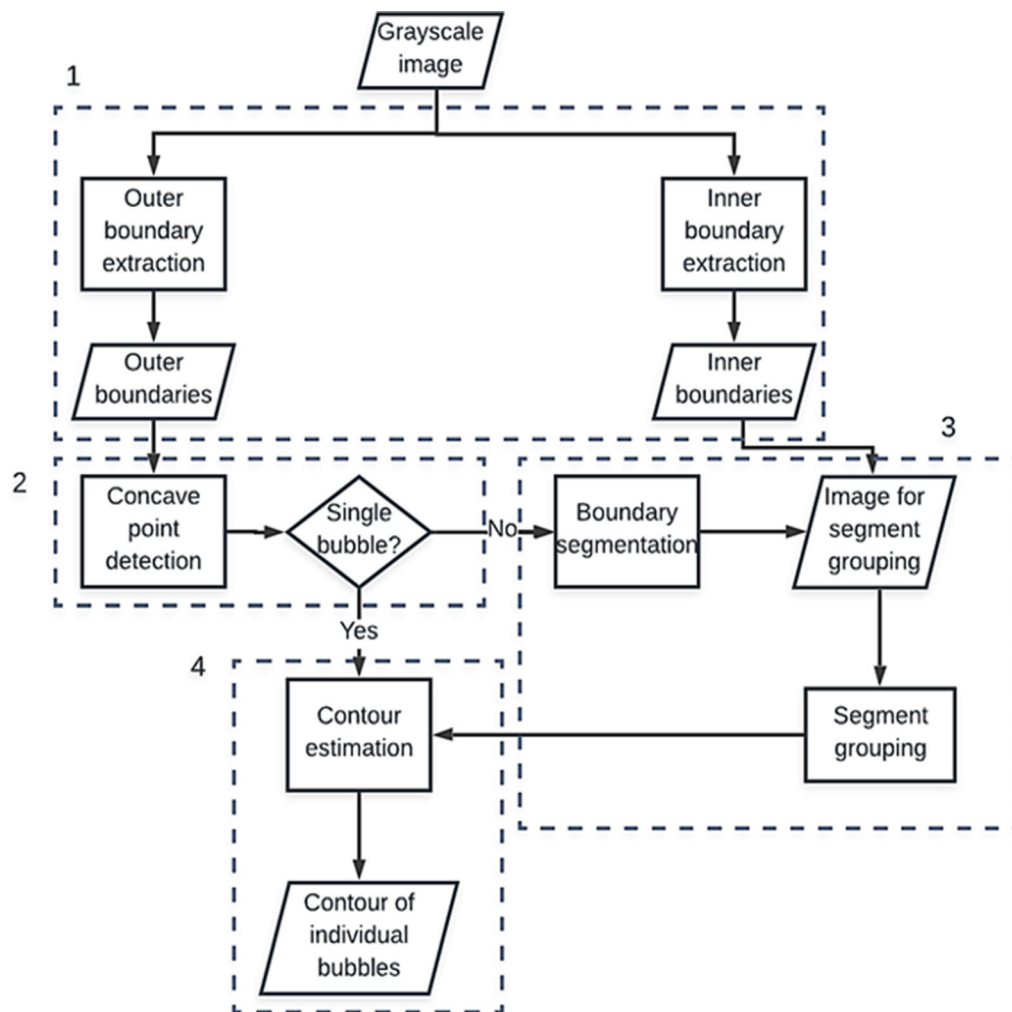


Fig. 4—Flowchart of the employed multistage digital image procedure for bubble detection.

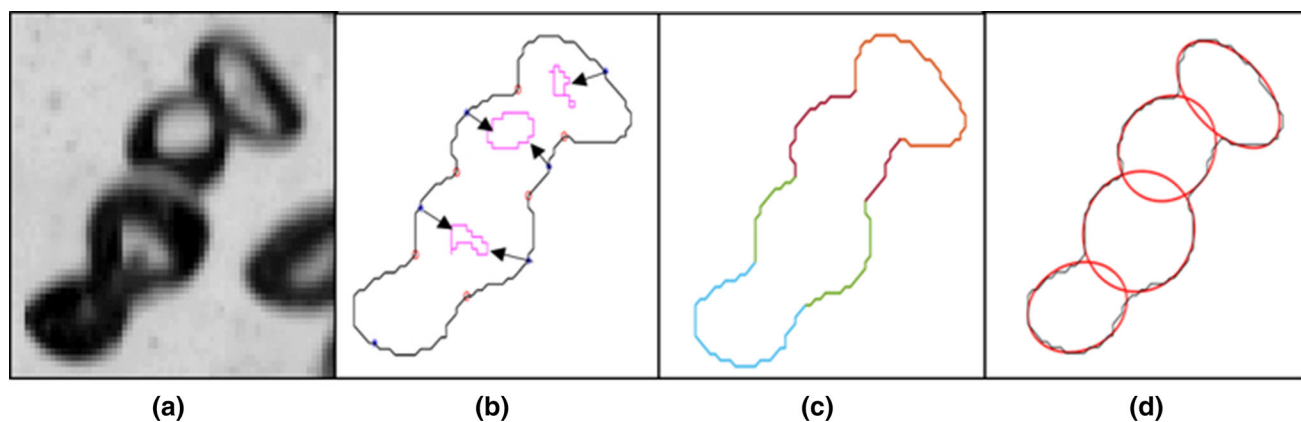


Fig. 5—(a) Segmentation of a bubble cluster that includes (b) boundary segmentation, (c) segment grouping, and (d) ellipse fitting.

rising profiles is 2.8 pct and the maximum deviation is 6.7 pct. The mean rising velocity deviates by 4.6 pct and the plume width by 5.2 pct.

In the process of data acquisition and processing, different sources of uncertainty have to be considered. During image acquisition, the lens distortion leads to a slight falsification of the results. This uncertainty

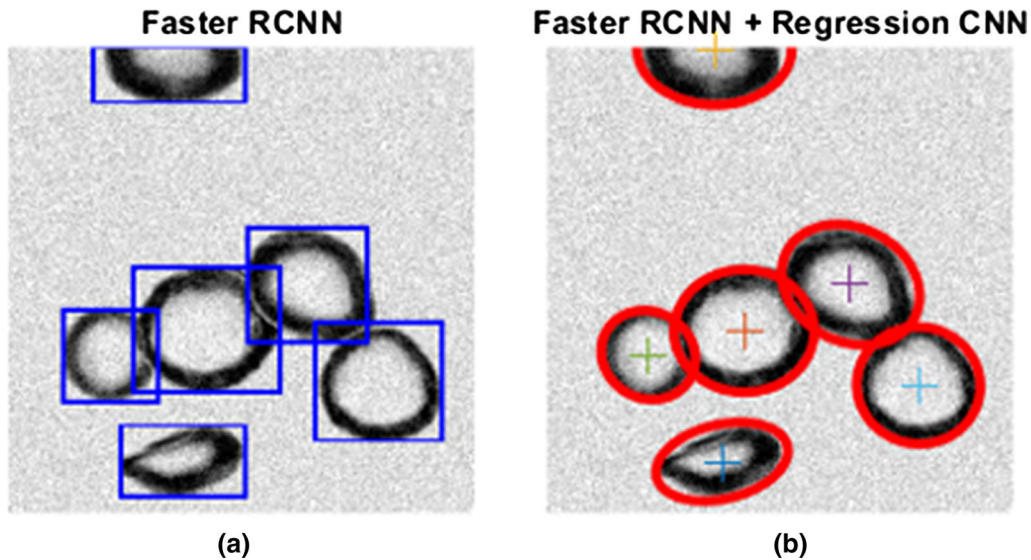


Fig. 6—BubCNN procedure. Bubble detection by (a) faster RCNN and (b) shape regression.

depends on the particular lens and calibration and can be quantified very accurately by the calibration procedure. Because a 60-mm lens is used, the uncertainty is below 1 pct. For each calibration, the exact value is given in the database. Another source of systematic uncertainty is the evaluation procedure. In Figure 7(a), the average bubble rising velocity in dependency of the radial distance from the plume center measured in the same test case is compared for the different detection methods. For swarm tracking, 500 bubbles are manually tracked. It can be seen that all methods yield very similar results. The largest deviation can be observed at the edges of the bubble column. Manual bubble clicking, in particular, shows the strongest deviation, while BubCNN and Imaging provide almost identical results. This can be explained by the fact that the number of manually clicked bubbles is relatively low, especially in the marginal areas. As a result, the values determined for clicking may not be statistically meaningful but are influenced by outliers. The mean relative deviation between BubCNN and Imaging is 2.5 pct, while it is about 5 pct between BubCNN and Clicking. The maximum deviation is 6 and 12 pct, respectively, at the periphery of the plume where the mean values are more affected by outliers. It should also be considered that the detection missing rate of automatic detectors depends on the local void fraction.^[63] Therefore, the measurement uncertainty might be slightly higher at lower heights because the local void fraction is highest close to the porous plug. The bubble size distributions obtained by the different techniques are shown in Figure 7(b). By comparison, it can be concluded that the systematic uncertainty introduced by the evaluation technique is about 4.5 pct.

A major problem is that only one high-speed camera is available. Thus, only two-dimensional data are available for a three-dimensional problem. Assumptions have to be made about the information in the third dimension. Regarding the bubble size, it is assumed that

bubbles are rotational symmetrically around the axis along the rising direction. With that, three-dimensional ellipsoids can be reconstructed from the ellipses. However, Fu and Liu^[65] showed that the bubble volume error of a single bubble is about 25 pct for a one-camera system, though it has to be clarified that this value strongly depends on the Reynolds and Eotvos numbers. The error can be significantly reduced in the case where a second camera is used. Similarly, the bubble rising profiles are biased because bubbles located at the image center are assumed to be in the plume center. However, the bubbles can have different locations on the unknown axis, actually being a significant distance from the plume center. Keeping that in mind, it is evident that the reconstructed profiles are flatter than the real ones. In addition, a second high-speed camera would be useful because three-dimensional rising tracks could be reconstructed as well, which would significantly increase the validation capabilities for bubble-related submodels. To the best of the authors' knowledge, a reconstruction with more cameras has not been made for a ladle yet.

It has been reported in the literature^[55] that water impurities have an impact on the bubble surface and, thus, on its shape and rising velocity. Therefore, one setting is repeated with tap water without tracer particles. It is found that the mean rising velocity is about 10 pct higher than in the equivalent experiment with tracers. In addition, the width of the plume is reduced by about 10 pct. This shows that the PIV tracers have an impact on the bubbles and probably also on the main flow. Further research is necessary to quantify this effect in more detail. For validation, this result implies that bubble drag models for impure systems, such as those of Tomiyama *et al.*,^[66] should be used.

During validation, it should also be considered that the bubble column fluctuates, flattening the averaged rising profiles in comparison to the instantaneous rising profiles, especially near the surface. Therefore, it is

Table IV. Advantages and Disadvantages of Different Bubble Detection Techniques

	Advantage	Disadvantage
Manual Evaluation	+ easy + low requirements on images + human intuition	– time consuming – not reproducible – no generalization
Conventional Image Processing	+ automated + reproducible + high accuracy + extensive statistical analysis	– less statistical evaluation – difficult to program – requires expert knowledge in image processing
BubCNN	+ automated + reproducible** + faster than conventional image processing on GPU + no expert knowledge in image processing required + extensive statistical analysis	– expensive equipment* – accuracy unknown a priori – expensive equipment*

*High-speed camera with a shutter speed of at least 1/1000 seconds and a strong, homogenous backlight.

**In case the same pretrained modules are employed.

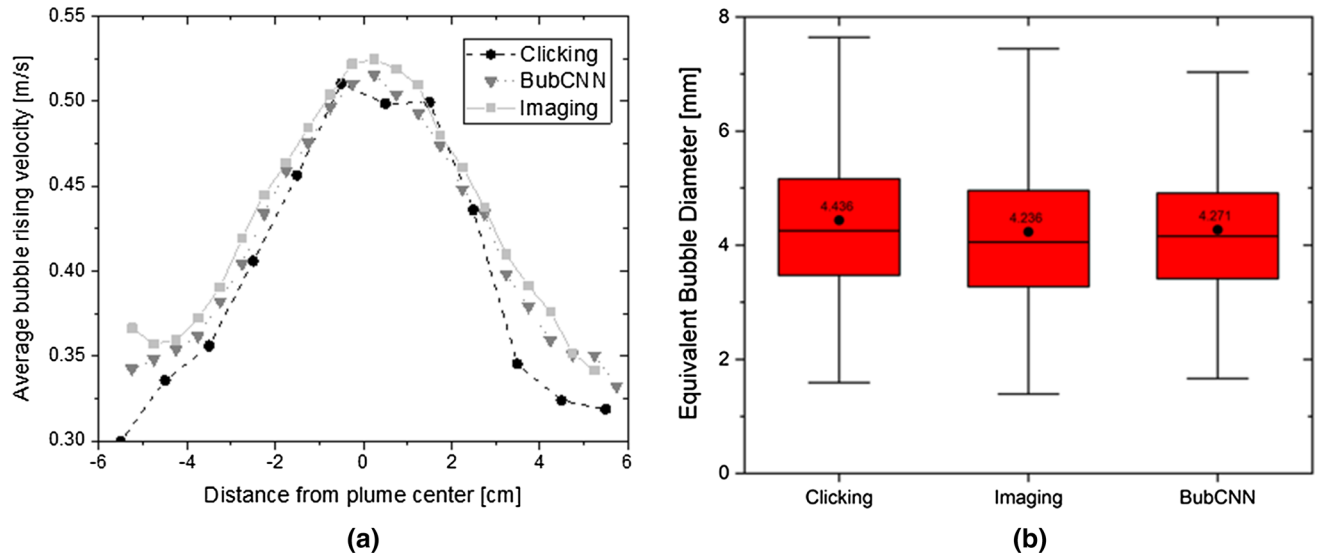


Fig. 7—Comparison of (a) the bubble rising profiles and (b) the bubble size distribution reconstructed with different detection methods.

recommended to either use the instantaneous profiles in the database or to use the same sampling time for an averaged profile.

In conclusion, the uncertainty of the bubble rising velocity profile, the mean bubble rising velocity, and the plume width can be estimated to be in a range of 10 pct. The main contributions to the total measurement uncertainty are the systematic uncertainty of the evaluation method and the aleatoric uncertainty. The impact on the latter could be reduced by averaging multiple measurements with the same settings.

The uncertainty of the bubble size measurements is about 25 pct for an individual bubble, but as this can be assumed to be a random uncertainty, it is reduced significantly by the large sample size in the range of 25,000. Thus, it can be concluded that the uncertainty

on the bubble size distribution is dominated by the systematic uncertainty introduced by the detection technique, which is about 5 pct.

For a final assessment of the uncertainty, Oberkamp and Smith^[48] proposed cross-validating the measurements by a different measurement technique such as electroresistivity probes. However, this technique is not available in the lab.

III. RESULTS

In Figure 8, example results for a flow rate of 2.4 slm and a height of 0.24 m are shown. Data for the other flow rates, heights, and more detailed results are provided in the database. The data can be used to

validate the plume region of the isothermal ladle flow as well as PBMs. The database comprises the following:

- (1) averaged rising velocity profiles (averaged over 10.9 seconds) (.xls and .mat format);
- (2) raw track data, including mean diameters, center positions over time, and tracked length (.mat format);
- (3) bubble size distributions (.xls format);
- (4) high-speed videos (shortened and compressed);
- (5) videos of full plume; and
- (6) additional documentation about the setup and procedure.

A. Single-Phase PIV

There are different nonintrusive flow measurement methods available to determine the flow field in a water model. Most prominent are LDA and PIV. For the given problem, PIV is more suitable than LDA because it conducts area measurements, capturing the velocity at multiple points. LDA, on the other hand, is a point measurement system, so for velocity measurements on a plane, a large number of measurement points need to be addressed. In addition, Deen *et al.*^[67] showed that LDA measurements need longer averaging times. On the other hand, LDA can be directly used in the plume region and can measure the velocity of both phases directly. In addition, planar PIV only provides two-dimensional information about volume-averaged local tracer displacements. Therefore, a highly resolved assessment of three-dimensional turbulence structures is not feasible. Nonetheless, the flow measurements are conducted by PIV in this work. Therefore, distinguishing between measurements out and into the multiphase plume region is necessary because the latter require additional preprocessing.

B. Experimental Procedure

The system is calibrated using the single-shot calibration system included in DaVis. For that, a calibration plate with an equidistant dot pattern with defined dot diameters and distances is placed in the measurement plane. In the calibration procedure, dots are automatically detected and the lens distortion and the magnification factor are computed based on the location of the detected dots. These factors are provided for each measurements setup individually in the database. The origin of the coordinate system is set at the center of the lower edge of the calibration plate and has to be marked manually.

Before each experiment, small bubbles that are attached on the model's walls are removed. After a new flow rate is adjusted or the gas purging is restarted, there is a wait time of 5 minutes so the flow can settle.

Single-phase PIV is applied in all regions outside the plume for flow rates of 1.2, 1.8, 2.4, and 3 slm. Measurements are made for the flow on the symmetry plane ($y = 0$), its perpendicular plane ($x = 0$), and closeups near the wall, near the free surface close to the

plume and two other locations given in Table V and shown in Figure 9. The purpose of the closeups is twofold. First, it provides more detailed data for locations of particular interest such as the shear layers at the surface or the wall. Second, it is used to estimate the measurement uncertainty of the velocity measurements. For each measurement, the averaged flow field and the velocity fluctuations are derived.

The PIV setup comprises a CCD double frame camera (ImageProX 4, resolution 2048×2048 pixel) and a double pulsed Nd:YAG laser (Litron LPU 550, $\lambda = 532$ nm) arranged as shown in Figure 10. Detailed specifications are provided in the database. A combination of two spherical lenses and one cylindrical divergence lens ($f = -10$ mm) widens the laser beam to a thin light sheet. The light sheet is positioned at different positions in the ladle model. The CCD camera is focused on the same plane with a small aperture. The water is seeded with tracer particles (Vestosint 1111, $\rho = 1.016$ g/cm³, $d = 50$ to 75 μ m) that follow the flow almost slip free.

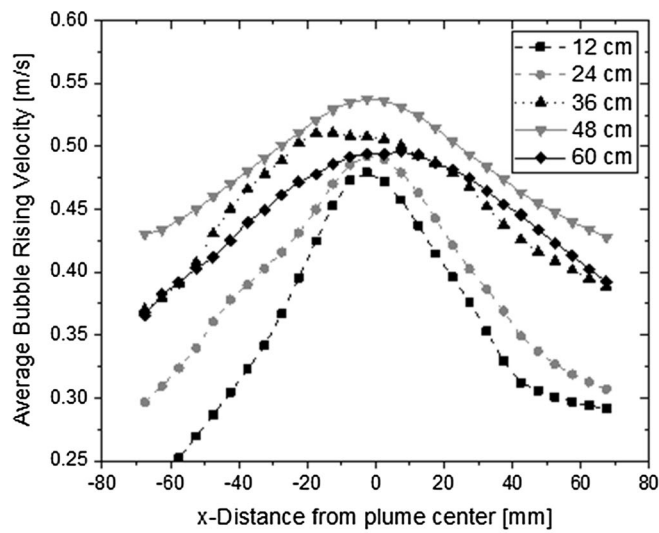
Two consecutive images with a delay of a few microseconds are taken, hereafter referred to as double frame. The laser is pulsed synchronously so that both frames are exposed sufficiently. The delay between the double frames is a trade-off so that the volume-averaged displacement of tracers $s(x,t)$ is sufficiently large but the out-of-plane displacements are still small. The value is provided for each experiment individually in the database.

Double frames are recorded with a rate of 5 Hz and first stored in the internal memory of the camera before they are transferred to the hard drive. Therefore, the maximum storable number of double frames is 75. In a previous work,^[50] it was shown that at least 1250 double frames should be used to derive significant mean flow fields. Thus, each measurement comprises 20 acquisition loops. That means, however, that the double frames are not in a temporally constant order. This does not affect either the mean flow field or the derived velocity fluctuation, but it prevents a frequency analysis of the flow.

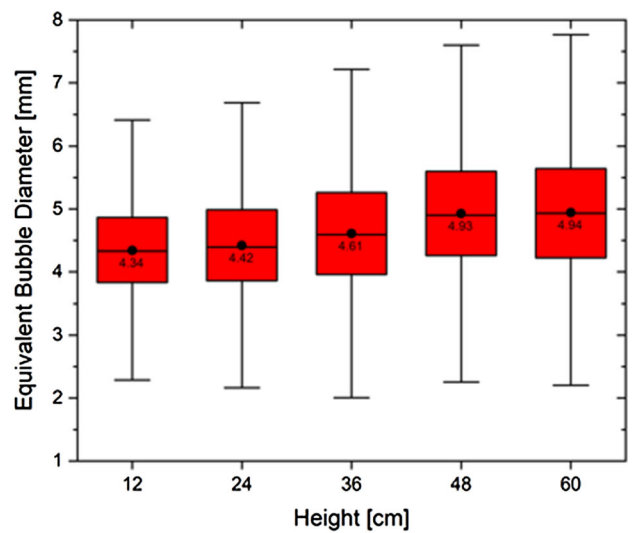
C. Data Processing

The double frames are processed, as shown in Figure 11. The parameters for each process stage can be found in the database.

To obtain the instantaneous vector fields, the double frames are analyzed by means of cross-correlation of the pixel intensity values. Thereby, the first frame is divided into interrogation areas. The interrogation areas are used as filters for a two-dimensional convolutional operation, which is the sum of an elementwise multiplication of the filter values and the light intensities of the receptive field. The receptive field is an area of the same size that is slid over the second frame. The convolutional operation yields its highest values in the case where the filter and the receptive field are similar. Thus, it detects the shift of particle patterns between the frames. To speed up the process, a Fourier transformation is



(a)



(b)

Fig. 8—(a) Bubble swarm rising velocity profiles and (b) equivalent bubble diameter distributions in dependency of the measurement height for a flow rate of 2.4 slm.

Table V. Coordinates of the PIV Closeups

	x (mm)	z (mm)	Δx (mm)
1	− 170	535	95
2	− 275	550	85
3	115	615	85
4	− 275	250	85

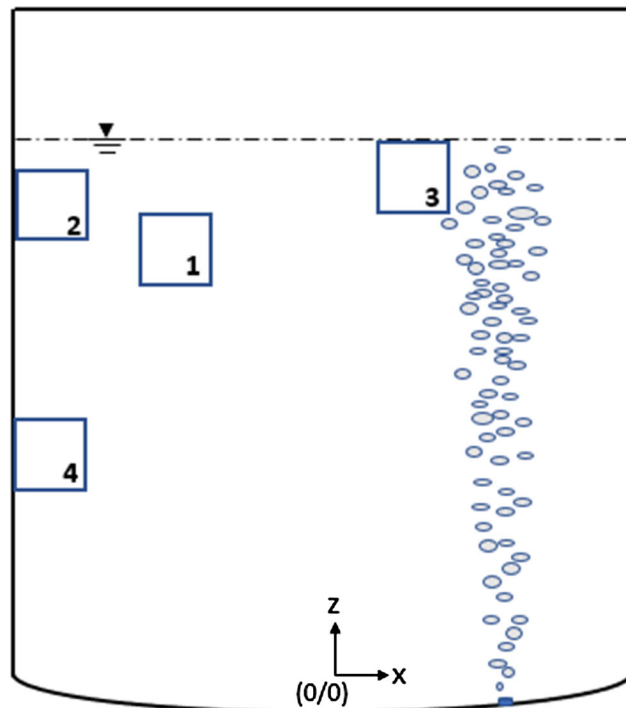


Fig. 9—Closeup locations.

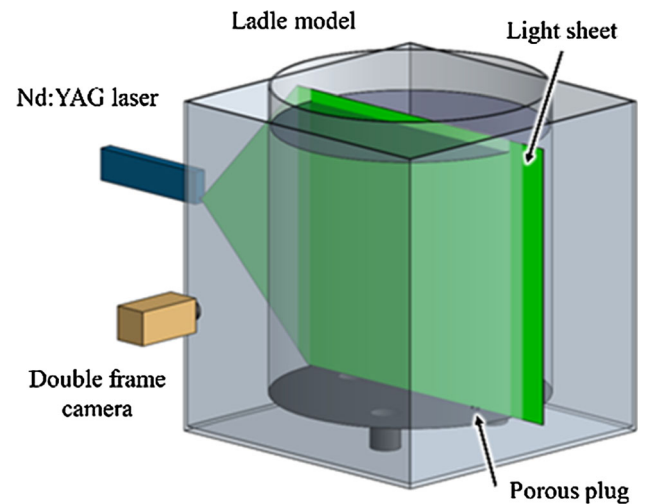


Fig. 10—Experimental setup PIV. Adapted with permission from Ref. [68].

applied to the filter as well as the receptive field. A volume-averaged velocity vector $\vec{v}(x, y, t)$ for each interrogation area can be computed by

$$\vec{v}(x, y, t) = \frac{\vec{s}(x, y, t) \cdot M}{\Delta t}, \quad [5]$$

where $s(x, y, t)$ is the volume-averaged displacement of a tracer pattern in the interrogation area, M is a conversion factor from pixels to a metric scale, and Δt is the delay between the double frames.

After vector postprocessing, the processed data sets are exported from DaVis and merged in MATLAB, which is used to derive the mean and the velocity fluctuation. The fluctuation of the measured velocity components is derived from the instantaneous profiles and the mean flow field by

$$u'_i = \text{RMSE } u_i = \sqrt{\frac{\sum_j^n (u_{i,j} - \bar{u}_i)^2}{n}}, \quad [6]$$

where RMSE u_i is the fluctuation of the velocity component i , \bar{u}_i is the mean velocity component, $u_{i,j}$ is the instantaneous velocity component, and n is the sample size.

D. Uncertainty Quantification

The measured values are affected by uncertainties introduced by different experimental settings or throughout the evaluation process. To quantify the effects, different measurements are conducted. To account for the aleatoric uncertainty, the same experiment is repeated 5 times, leaving all parameters and the setup unchanged but stopping (5 minutes) and restarting the gas flow rate. It is found that the results varied slightly, with a mean relative standard deviation for the flow on the symmetry plane of 5.6 pct. Random, epistemic uncertainty can be reduced by the number of double frames included to derive the means and velocity fluctuation. With 1500 double frames used, it is found that the last 50 double frames change the mean value by less than 1.5 pct. More critical and more difficult to quantify are systematic, epistemic uncertainties, which can be caused by the scaling factor M , uncorrected lens distortion, and the definition of the coordinate system's origin. To account for that, the experimental setup is changed by shifting both the camera and the laser. It is found that the mean relative deviation to the mean flow field obtained when the old calibration is 10.1 pct. While the relative deviation is merely about 5 pct at the main flow, the highest deviations of up to 200 pct are found at locations of high velocity gradients, particularly close to the free surface and the toroid. In addition, the deviation is undirected. The velocity magnitude is higher in some areas and lower in others. Because of that, it can be assumed that the manual definition of the coordinate system origin is the main source of uncertainty. Although the origin's positions differ only in a few

millimeters, this can have a significant effect on the measured flow components in areas of high velocity gradients.

Another source of uncertainty is the delay between the recordings that form a double frame. The delay affects the measurements in different ways. It determines the impact of out-of-plane displacements and the scaling effect. Moreover, the derived vectors are filtered in the spatial as well as in the temporal domain. Because of that, the smallest vortex structures cannot be resolved, which might cause a small underestimation of the velocity fluctuation. In addition, PIV assumes a linear unaccelerated movement of the tracers between a double frame. Because of that, the calculated flow velocity can be slightly smaller than the actual one, an effect that might be more pronounced for longer delays. To estimate the uncertainty introduced by the delay, measurements are made with a shorter (15.000 μs) and a longer (25.000 μs) delay and compared to the regular delay (20.000 μs) but the same calibration. It is found that a shorter delay deviates from the regular delay by 6.4 pct, on average, while a longer delay deviates 5.6 pct, which is close to the aleatoric uncertainty. The sum of the deviations for both cases is almost zero, so it can be concluded that there is no directed, systematic uncertainty introduced by the delay, neither on the mean, nor on the velocity fluctuation, as long as it is not too short or too long.

Like the delay, the resolution of the measurements can have an impact on the results. A reason is that the local tracer displacement is volume averaged. Thus, it can be assumed that smaller interrogation areas with a higher resolution yield more accurate results. On the other hand, the number of tracers per interrogation area decreases, which might increase the inaccuracy of cross-correlation. In addition, the resolution affects the smallest resolvable vortex structures and might have a decisive impact on the fluctuation velocity. To analyze these effects, the measurements of the full symmetry plane are compared with the closeups. It is found that the mean relative deviation of the velocity component is 13.3 pct, while it is 41.1 pct for the velocity fluctuations. However, some values exceed those of the full plane, while others are below it. Thus, it can be concluded that the deviation is not due to a systematic uncertainty caused by the aforementioned reasons but by an uncertainty caused by the calibration. Unfortunately, all closeups are at locations of high velocity gradients. Because of the calibration uncertainty, the derived values of the closeup measurements should not be used

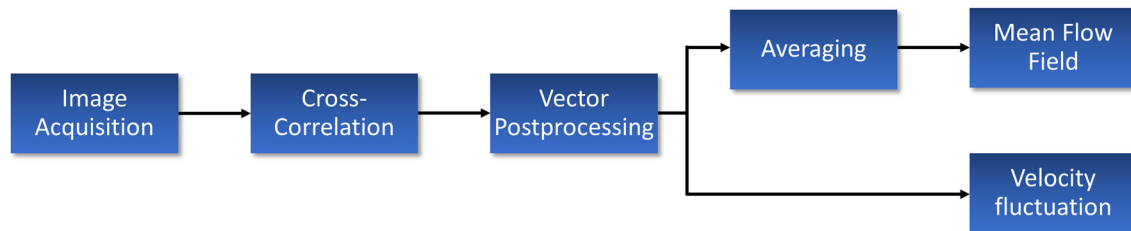


Fig. 11—Flowchart single-phase PIV.

for a quantitative validation. However, they still provide quantitative insights on areas of particular interest, such as shear layers, with a higher resolution than the measurements on the full plane.

A crucial assumption of PIV is that the tracer follows the flow slip free. Although PIV is generally considered a nonintrusive measurement technique, it is found in the uncertainty assessment of the plume analysis that the tracers have an impact on the bubble rising velocity. Since it is not possible to conduct PIV without tracers, an effect on the main flow is difficult to quantify. However, the measurements are repeated with fluorescent rhodamine-B coated PMMA particles ($\rho = 1.050 \text{ g/cm}^3$, $d = 50 \text{ to } 100 \text{ }\mu\text{m}$) to estimate the effect of the choice of tracers. For that measurement, a cut-off filter is attached to the lens, so the system has to be recalibrated. The mean relative deviation from the mean flow field is found to be 6.9 pct, which is slightly above the relative standard deviation but below the uncertainty introduced by a new calibration. Thus, it can be assumed that the uncertainty by the slip between the tracer and fluid is comparatively small. However, additional measurements with other techniques are necessary for a final conclusion on that topic.

Cross-correlation yields some incorrect tracer shifts. To correct these errors, different vector validation methods are available; most prominent are median filters and velocity component constrains for the spatial domain and standard deviation filters for the temporal domain. A median filter computes a median vector of an $n \times n$ grid around each vector. A vector gets rejected in the case in which its deviation from the median vector exceeds a predefined threshold. Velocity component constrains restrict all vectors to a specified range. Vectors exceeding the range with one component will be rejected. A standard deviation filter compares the instantaneous vector with the spatial mean. If the instantaneous vector differs more than a multiple of the standard deviation, the instantaneous vector gets rejected. Rejected vectors will be replaced by the average of all nonzero neighbors. The correct choice of methods and parameters can be challenging. In the case where postprocessing is too weak, the results are affected by the wrong data. In the case where it is too strong, correct signals get discarded. The uncertainty introduced by vector postprocessing is estimated by using different methods and combinations of these for the same data set. It is found that a standard deviation filter that removes all signals that are 1.5σ from the mean is insufficient for postprocessing. For the other investigated methods, a mean flow field is computed and the results of the different vector validation methods are compared to the mean. In the case where only a median filter with a filter width of 7 is used, the mean relative deviation is about 4 pct. With velocity constrains, it deviates by 1.2 pct. By a combination of the different methods, the deviation decreases to 0.8 pct. In conclusion, the uncertainty introduced by vector postprocessing is comparably low, in a range of about 2 pct, in the

case where the methods and parameters are chosen correctly. An improper choice of vector postprocessing can increase the uncertainty to about 25 pct.

Among commercial PIV software, there are some open-source projects as well. Here, PIVlab^[69] is compared to DaVis. For that, all acquired double frames for a setup are exported from DaVis and processed in PIVlab with the same parameters. The mean relative deviation is about 0.5 pct. The maximum deviation is about 3 pct. The velocity magnitude obtained with PIVlab is slightly lower close to the walls, while it is slightly higher in the ladle center. This indicates the influence of a barrel lens distortion on the results. In DaVis, single-shot calibration is used to automatically correct the image distortion, while in PIVlab, that would be a separate preprocessing step that is not taken in this work. However, it is likely that an appropriate image calibration, for instance, with the MATLAB Camera Calibration App, would further minimize the deviation between the results. So, both programs can be used but have some specific advantages. DaVis can also be used for calibration, image acquisition, laser-camera synchronization, and further control steps. That makes the use of PIV relatively simple and beginner friendly. On the other hand, the software license fee is expensive, while PIVlab is an open-source project. In addition, according to the authors' opinion, customization of the pre- and postprocessing is easier.

In addition to the hitherto discussed uncertainty, one has to keep in mind that planar PIV only measures two of the three velocity components. For a comparison of the flow profiles, one should include the two measured components of the numerical model as well. The same applies for a comparison of the velocity fluctuation components. Any conclusion on the third component, including the derivation of the TKE, may lead to a significant increase in uncertainty.

In summary, the mean overall uncertainty of the PIV measurements is about 10 pct, mainly caused by the calibration procedure. However, the uncertainty can increase up to 200 pct at locations of high velocity gradients. Although the calibration uncertainty is systematic when only two different calibrations are compared, it is reasonable to assume that it is random when a larger number of different experimental settings and calibrations are used. By that, the measurement uncertainty can be further reduced. However, to the authors' knowledge, no study is known in which this strategy was employed. The main reason is that the velocity to coordinate mapping is only necessary for a quantitative validation. For deterministic models or a quantitative comparison, the deviation, which does not affect the overall flow structure, is negligible. For the current accuracy level of numerical models, the deviation level is too small to justify the increased experimental effort. However, in the future, experiments with a lower uncertainty might be required for a fine-tuning of the numerical models.

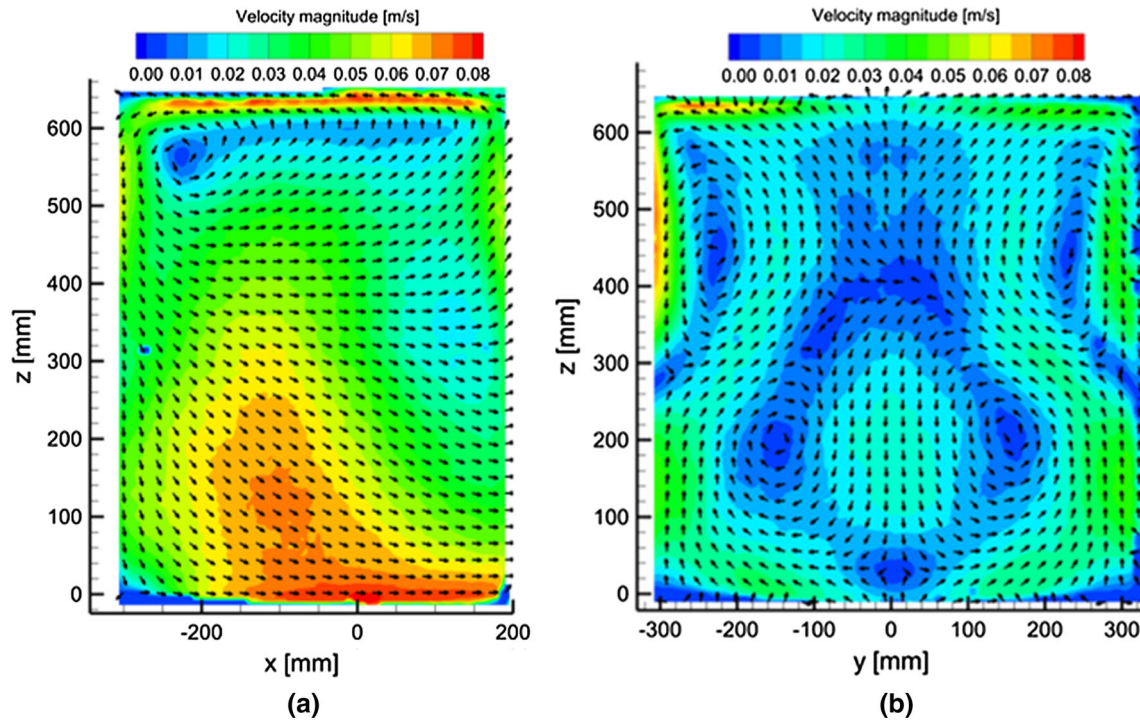


Fig. 12—Velocity magnitude profiles for (a) the symmetry plane and (b) the perpendicular plane.

IV. RESULTS

In Figure 12, example results for the symmetric plane and its perpendicular plane for a flow rate of 2.4 slm are shown. Data for the other flow rates, closeups, and a more detailed documentary are provided in the database:

- (1) averaged flow profiles (.dat, .xls, and .mat format);
- (2) profiles of velocity fluctuations (.dat, .xls, and .mat format);
- (3) instantaneous profiles (.mat format);
- (4) sample images; and
- (5) additional documentation about the setup and procedure.

A. Multiphase PIV

Measurements in the plume region are conducted at five different heights ($z = 0.12, 0.24, 0.36, 0.48$, and 0.60 m) for four different flow rates (1.2, 1.8, 2.4, and 3.0 slm). For all measurements, the averaged velocity profile and the velocity fluctuation are derived.

In contrast to single-phase PIV, measurements in the bubble swarm region require special precautions. A reason is the existence of phase boundaries that reflect some of the illumination of the light sheet, resulting in shadow regions behind the bubbles and very intense light signals at the reflection area. Because of that, PIV measurements in bubble plumes are limited to medium void fractions of about 5 pct.^[55] A major challenge for multiphase PIV is to distinguish the signals generated by

the tracers from those generated by the phase boundary. Because the average shift is computed for each interrogation area, an evaluation of that mixed signal would overestimate the actual flow velocity. Different discrimination techniques have been proposed in the literature, which are discussed by Brücker^[55] and Deen *et al.*^[70]

Nowadays, the most common method is to use fluorescent particles and a cutoff filter, as illustrated in Figure 13. Deen *et al.*^[67] referred to that technique as PIV/laser-induced fluorescence. Usually, rhodamine-B dotted particles with a stimulation maximum at a wavelength of 540 nm are used. The emission maximum is at a wavelength of 584 nm, so cutoff filters can be used to distinguish between reflections and fluorescence.^[55]

For the multiphase PIV measurements, a similar experimental setup as for single-phase PIV is used. However, fluorescent rhodamine-B coated PMMA particles ($\rho = 1.050$ g/cm³, $d = 50$ to 100 μ m) are used as tracers and the CCD camera is equipped with a cutoff filter (>540 nm) and a 60-mm lens. Because the velocity in the plume region is much higher than in the single-phase region, the delay between the double frames Δt is reduced to 1500 μ s.

For each measurement setup, 10 loops of 75 double frames are acquired using DaVis. The mean plume profiles and velocity fluctuation are derived from those images by the procedure shown in Figure 14. For image preprocessing, MATLAB is used, while cross-correlation and vector postprocessing are made with PIVlab. The images are calibrated by the same procedure employed for bubble swarm tracking.

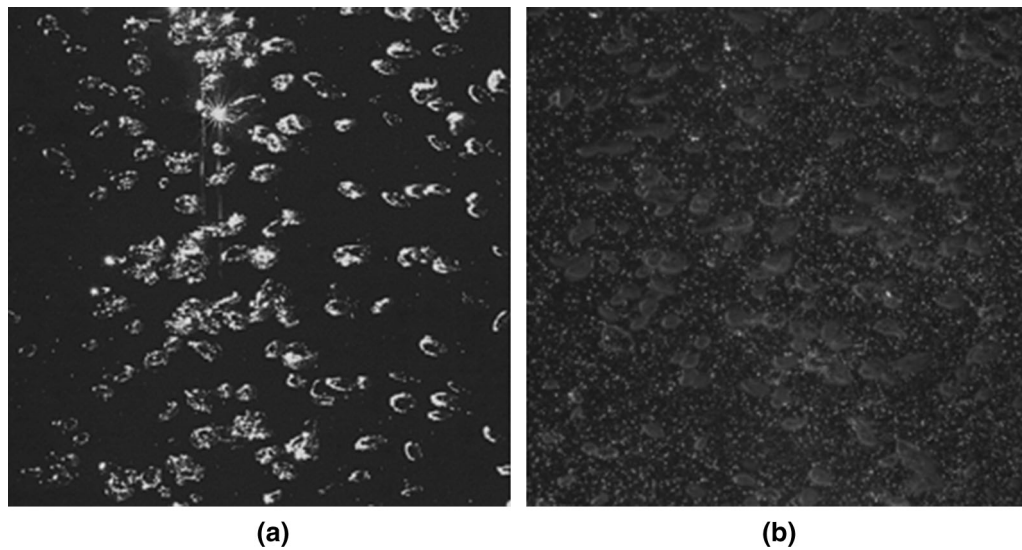


Fig. 13—Multiphase PIV images (a) without and (b) with cutoff filter.

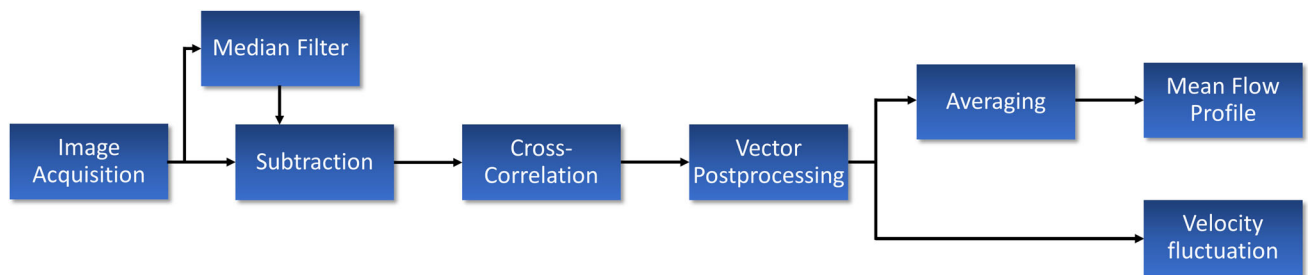


Fig. 14—Flowchart multiphase PIV.

As discussed in the literature^[55] and shown by bubble swarm tracking, the tracers have an impact on the flow. The effect is most noticeable in the plume region as tracer particles can accumulate on the phase boundaries. Because of that tracer accumulation, additional post-processing becomes necessary. Similar to Deen *et al.*,^[71] a median filter with a kernel size of 7×7 is employed, as shown in Figure 15. Thereby, small sized signals get replaced by a median intensity value. The larger signals generated by the bubbles remain (a). Afterward, that image is subtracted from the original one so that only the tracer signals remain (b).

For comparison, the average profiles were zero centered so that the location of maximum velocity is defined as the plume center. For validation, the bell-shaped rising profiles are described by their height and width. The height is defined as the maximum velocity and the width as the distance of the points around the maximum at which the speed was halved. For a more precise determination of the distance, the profiles are interpolated linearly on a scale from -0.1 to 0.1 m with a width of 0.1 mm.

B. Uncertainty Quantification

For the estimation of the PIV measurements, uncertainty in the plume region, similar considerations as for single-phase PIV apply. However, there are some specifics to be considered. The measurement points are not assigned to physical coordinates, but the position of the maximum velocity is assumed to be the center of the plume. This minimizes the uncertainty caused by the definition of the origin, which is decisive for the uncertainty of the single-phase PIV measurements. The remaining uncertainty due to an incorrectly marked height in the image is found to be very small. Even if the height is over- or underestimated by 10 mm, the deviation of the maximum velocity is found to be 2.5 pct. Furthermore, a 60 -mm rather than a 32 -mm lens is used to increase the accuracy of the scaling factor and decrease the potential uncertainty by improperly corrected lens distortion. The systematic uncertainty of the scaling factor is found to be below 1 pct. The residual uncertainty, defined as the means change by the last 50 frames, is found to be about 0.5 pct.

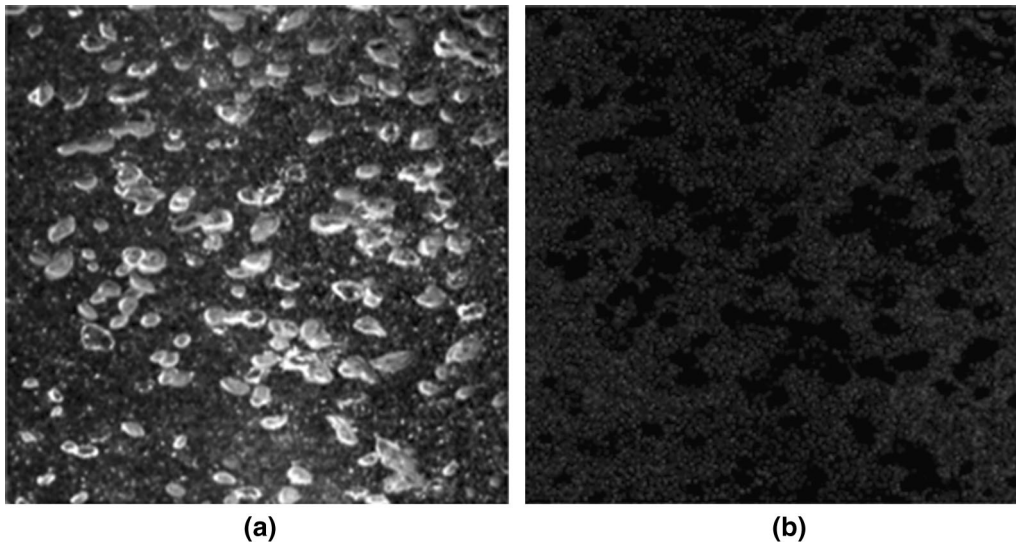


Fig. 15—(a) Multiphase PIV image with median filter and (b) subtraction of (a) from the original image.

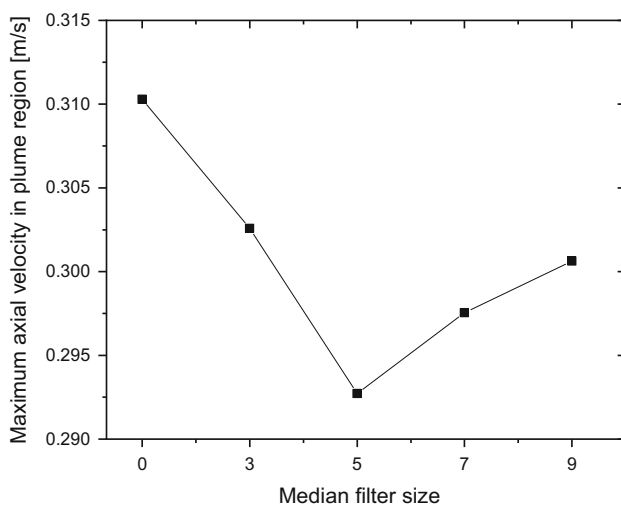


Fig. 16—Maximum axial velocity derived by multiphase PIV in dependency of the median filter size.

In comparison to single-phase PIV, an additional systematic uncertainty caused by image preprocessing must be considered. Brücker^[55] reported that the filter width of the median filter has an impact on the results. To quantify this uncertainty, measurement with the same parameters is preprocessed with different median filters. The maximum velocity in dependency of the filter width is shown in Figure 16. As expected, a direct use of the image (filter width = 0) yields the highest maximum velocity. However, as reported in the literature,^[67] the velocity is increased by vectors generated by bubbles or tracers attached to the bubble. With a filter width of 3, too many tracers get removed. The velocity profile becomes unsteady rather than resulting in a smooth bell shape. A filter width of 5 or 7 seems to be an ideal

choice. However, it should be remembered that the optimal filter width depends on the resolution and the bubble and tracer size. In the case of a filter width of 9, the bubbles are not fully removed. Thus, the maximum velocity increases. However, the uncertainty introduced by the filter width is relatively small. If properly chosen, it is found to be on the order of 1 pct for the profile's width as well as the maximum velocity.

The systematic uncertainty by the tracers, which is known to have an effect on bubble velocity, shape, and coagulation and breakup behavior,^[55] is very difficult to quantify. Follow-up research is required, for instance, by a cross-validation by other measurement techniques. In conclusion, the overall uncertainty for the multiphase PIV measurements is found to be about 5 pct.

V. RESULTS

Exemplary results are shown in Figure 17 for a flow rate of 2.4 slm. Additional data can be found in the database:

- (1) averaged flow profiles (.dat, .xls and .mat format);
- (2) velocity fluctuation profiles (.dat, .xls and .mat format);
- (3) instantaneous profiles(.txt format);
- (4) sample images; and
- (5) additional documentation about the setup and procedure.

A. Validation Guidelines

A decisive feature of a strong-sense database is the standardized validation process.^[49] To the best of the authors' knowledge, a standard validation procedure does not exist in the field of metallurgy hitherto. Here, some guidelines are developed for the usage of the provided data and a validation procedure for isothermal flows is proposed. Key features are identified, discussed,

and summarized of a validation score that ranges from 0 to 100, rather than a complex validation system. A decision was made to derive a single number accuracy assessment criterion because it allows a direct comparison of the numerical models and is, therefore, beneficial for a structured optimization of the models. However, it brings the difficulty of capturing all information in a single number without missing relevant data. For example, in a former work,^[50] it is shown that an evaluation of the flow structure by using line plots rather than contour plots can cause misleading results. Therefore, flow characteristics that capture all essential information have to be determined. For the given problem of an isothermal flow, the choice was made to use five different subsystems: the location of the toroid, the velocity components, the velocity fluctuations, averaged over the entire symmetry plane and on eight monitoring points, and a characterization of the plume area for both the bubbles and the fluid. Standardized evaluation procedures and metrics are described in detail subsequently for all criteria. Following these instructions results in a scoring for each subsystem. The final score is the sum of scores of all subsystems. That strategy allows both an overall optimization based on a single number and the detection of optimization potentials by subsystem scorings. A difficulty is that the score has to be “fair” and sensitive. Both properties rely on a suitable balance of scoring ranges of the subsystems. It appears that the proposed ranges can only be a first draft and that a standardized procedure can only be established over time through exchange with users. Programs are provided in the database that allow an automated computation of the validation scores for a flow rate of 2.4 slm for Reynolds-averaged turbulence models and time-sampled LES models. For its usage, three sources of data need to be exported from the CFD solver in a specified order:

- (1) flow on the symmetry plane ($y = 0$) as columnwise ASCII code: x -coordinate, y -coordinate, z -coordinate, RMSE z -velocity, RMSE x -velocity, mean- z -velocity, and mean- x -velocity;
- (2) flow on the perpendicular plane ($x = 0$) as columnwise ASCII code: x -coordinate, y -coordinate, z -coordinate, RMSE z -velocity, RMSE y -velocity, mean- z -velocity, mean- y -velocity; and
- (3) bubble characteristics as columnwise csv files, saved every 0.1 seconds for 10.9 seconds: x -coordinate, y -coordinate, z -coordinate of bubble center, and z -velocity.

For Reynolds-averaged turbulence models, the TKE rather than the RMSE velocity components have to be exported. Text menu commands and a user-defined function to export these values from Ansys Fluent are provided in the database.

To make the results independent of the numerical mesh, the CFD results are interpolated on an equidistant grid with a spacing of 5 mm, ranging from $z = 5$ to 640 mm and $x = -310$ to 310 mm for the symmetry plane and $y = -310$ to 310 mm for the perpendicular plane. The chosen interpolation method is the triangulation-based linear interpolation.

The validation score s_i for each subsystem ranges from $s_{i,\max}$ to 0 and is computed by

$$s_i = \min \left(\max \left(s_{i,\max} - \frac{s_{i,\max} \cdot (|v_{\text{measured}} - v_{\text{CFD}}| - u_i)}{r_i}, 0 \right), s_{i,\max} \right), \quad [7]$$

where u_i is the relative measurement uncertainty, v_{measured} is the measured value, v_{CFD} is the computed value, and r_i is the scoring range. Independent of the subsystem, the deviation can be expressed as either relative or absolute.

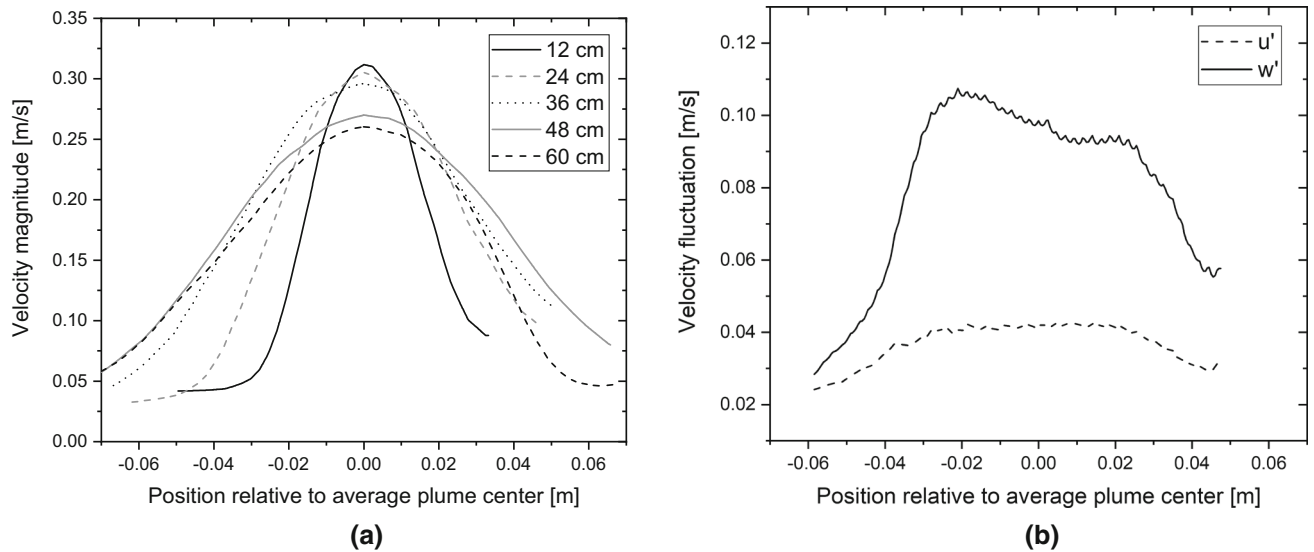


Fig. 17—Plume velocity magnitude profiles in dependency of the measurement height for (a) a flow rate of 2.4 slm and velocity fluctuation at 24 cm and (b) a flow rate of 2.4 slm.

B. Toroid Location

The first assessment criterion is the location of the toroid on the symmetry plane. The toroid is defined as the location at which the flow direction changes most pronouncedly to form a circular flow structure. Because of that, the toroid is not crossed by any streamlines. The toroid location is an indicator of the general flow structure and death zones, which are essential for mixing and temperature homogenization in the ladle process. Experimentally, the toroid $\vec{\text{tor}}_{\text{measured}}$ is found at $x = -222$ mm, $z = 564$ mm by PIV measurements on the full symmetry plane. The measurement uncertainty is 10 mm. With a maximum validation score of 10 and a maximum toroid distance range of 400 mm, Eq. [7] becomes

$$s_{\text{toroid}} = \min \left(\max \left(10 - \frac{10 \cdot (|\vec{\text{tor}}_{\text{measured}} - \vec{\text{tor}}_{\text{CFD}}| - 10)}{400}, 0 \right), 10 \right), \quad [8]$$

where $\vec{\text{tor}}_{\text{CFD}}$ is the computed toroid location. The automated detection of the toroid location can be difficult. Using the minimum of velocity magnitude on the symmetry plane can be misleading. Instead, the direction of the velocity vectors is used as a criterion. The velocity vector direction is computed by

$$\tan(\varphi(x, z)) = \frac{u(x, z)}{w(x, z)}, \quad [9]$$

where φ is the velocity direction in degree, ranging from -180 to 180 deg. The toroid is detected by applying a two-dimensional convolution on the absolute velocity direction field using the 5×5 filter f . The best results are obtained with

$$f = \begin{bmatrix} 1 & 0 & 0 & 0 & -1 \\ 2 & 1 & 0 & -1 & -2 \\ 5 & 2 & 0 & -2 & -5 \\ 2 & 1 & 0 & -1 & -2 \\ 1 & 0 & 0 & 0 & -1 \end{bmatrix}. \quad [10]$$

C. Averaged Velocity

The second assessment criterion is a comparison of the averaged mean velocity components and their fluctuation on the symmetry plane in the area of $x [-320:200]$, $z = [0,640]$. These factors indicate whether the numerical model is generally suitable to correctly predict temperature homogenization, mixing, and inclusion agglomeration. The mean values are summarized in Table VI. The measurement uncertainty is found to be below 0.001.

Each component can achieve a maximum scoring of 5. The scoring range r_i is set to 0.02 m/s. Thus, Eq. [7] becomes

$$s_{\text{means}} = \min \left(\max \left(5 - \frac{5 \cdot (|v_{\text{measured}} - v_{\text{CFD}}| - 0.001)}{0.02}, 0 \right), 5 \right). \quad [11]$$

For time-sampled LES models, both velocity fluctuation components are directly accessible. For RANS models, which are based on the assumption of isotropic turbulence, the velocity fluctuations are derived from the turbulent kinetic k energy by

$$u' = w' = \sqrt{\frac{2}{3}k}. \quad [12]$$

D. Monitoring Points

Using the mean values for an accuracy assessment is a good measure to evaluate the general ability to depict mixing, homogenization, and inclusion agglomeration. On the other hand, the mean values miss most details. Thus, the third assessment criterion is the flow values at monitoring points. Six monitoring points are placed on the symmetric plane and two additional ones on the plane perpendicular to it. For each monitoring point, the two measured velocity components, the direction of the velocity vector and the two measured velocity fluctuations, are evaluated with a maximum score of 1 for each component. The flow values are derived from PIV measurements of the full planes. They are summarized in Table VII.

It is important that only the actual measured velocity components are included in the validation procedure. The measurement uncertainty u_i is found to be below 0.0025 m/s for velocity components and 0.0015 m/s for velocity fluctuation. The scoring range r_i is set to 0.02 m/s for velocity components and 0.005 m/s for the velocity fluctuation.

$$s_{\text{monitor}} = \min \left(\max \left(1 - \frac{1 \cdot (|v_{\text{measured}} - v_{\text{CFD}}| - u_i)}{r_i}, 0 \right), 1 \right). \quad [13]$$

The deviation in the velocity vector direction is computed by

$$\cos \theta = \frac{\vec{m}_{\text{measured}} \times \vec{m}_{\text{CFD}}}{|\vec{m}_{\text{measured}}| \cdot |\vec{m}_{\text{CFD}}|} s, \quad [14]$$

where \vec{m} is the velocity vector. With a scoring range of 45 deg and a measurement uncertainty of 7 °C, the scoring is computed by

$$s_{\text{monitor,angle}} = \min \left(\max \left(1 - \frac{1 \cdot (\theta - 7^\circ)}{45^\circ}, 0 \right), 1 \right). \quad [15]$$

Table VI. Averaged Mean Velocity and Velocity Fluctuations on the Symmetry Plane

	u	u'	w	w'
Magnitude (m/s)	0.0289	0.0202	− 0.0118	0.0194

Table VII. Flow Components for the Monitoring Points

Point	<i>x</i> (mm)	<i>y</i> (mm)	<i>z</i> (mm)	<i>u</i> (m/s)	<i>u'</i> (m/s)	<i>v</i> (m/s)	<i>v'</i> (m/s)	<i>w</i> (m/s)	<i>w'</i> (m/s)
1	− 200	0	400	0.0398	0.0189	—	—	− 0.0120	0.0241
2	− 200	0	200	0.0426	0.0174	—	—	− 0.0439	0.0171
3	− 100	0	250	0.0549	0.0197	—	—	− 0.0323	0.0160
4	0	0	100	0.0573	0.0191	—	—	− 0.0272	0.0141
5	0	0	100	0.0407	0.0237	—	—	− 0.0139	0.0162
6	150	0	500	0.0134	0.0180	—	—	0.0147	0.0152
7	0	0	200	—	—	− 0.0007	0.0195	− 0.0228	0.0153
8	0	150	450	—	—	0.0012	0.0181	0.0215	0.0185

E. Plume Velocity

The fourth assessment criterion is the velocity of the bulk liquid in the plume zone. In the ladle, this region is characterized by the highest mixing energy, velocity, and velocity fluctuation. It is important for inclusion removal, slag eye formation, and PBMs. As shown in a previous work,^[50] a detailed knowledge about the plume region is also crucial for the improvement of all bubble-related multiphase models. For validation, the averaged velocity magnitude profiles at five different heights are used. The profiles are characterized by height and width, as defined previously, which both can have a scoring of up to 1.5 for each height. The measurement uncertainty is found to be 5 pct. The scoring range for both features is set to 50 pct. By that, the scoring can be computed by

$$s_{\text{plume}} = \min \left(\max \left(1.5 - \frac{1.5}{0.5} \cdot \left(\frac{|v_{\text{measured}} - v_{\text{CFD}}|}{v_{\text{measured}}} - 0.05 \right), 0 \right), 1.5 \right). \quad [16]$$

The numerical profiles are extracted from the interpolated symmetry plane, maximum centered and interpolated on a finer scale like the measured data.

F. Bubble Rising Velocity

The final assessment criterion is the bubble rising velocity. Like the plume velocity, it is important for the improvement of all bubble-related submodels. Furthermore, the bubble rising velocity is a crucial factor for inclusion removal. The procedure described here is derived for the Lagrangian discrete particle model. For validation, the mean bubble rising velocity and the width of the plume, as defined previously, are used. Both can receive a maximum scoring of 1.5 for each height. The measurement uncertainty is 10 pct and the scoring range is set to a maximum deviation of 100 pct:

$$s_{\text{bubble}} = \min \left(\max \left(1.5 - \frac{1.5}{0.5} \cdot \left(\frac{|v_{\text{measured}} - v_{\text{CFD}}|}{v_{\text{measured}}} - 0.1 \right), 0 \right), 1.5 \right). \quad [17]$$

The numerical bubble rising velocity is computed by importing all bubble characteristics in the exported csv-files. Bubbles are used in the case where their *z*-position is the specified height ± 0.05 m. Note that the bubble rising velocity is defined as the *z*-velocity.

Table VIII. Employed Submodels for the Example Numerical Models

Subsystem	Model
Turbulence	standard <i>k</i> – ϵ model/LES (Germano subgrid model)
Free Surface	volume of fluid model
Discrete	Lagrangian discrete
Bubbles	phase model
Bubble Size	Gaussian distribution
Considered	buoyancy
Forces	drag (Tomiyama) lift (Tomiyama) virtual mass
Time-Step	0.005 s
Size	
Mesh	hexahedral ($\Delta x \sim 6$ mm)

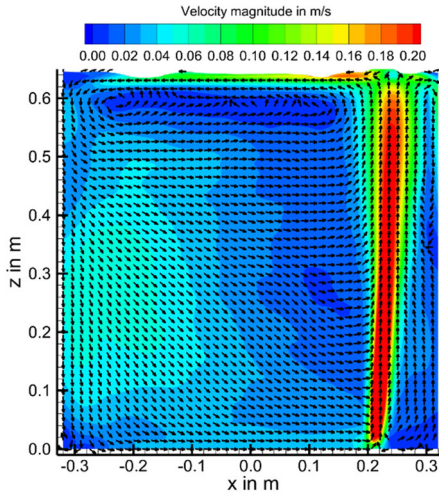
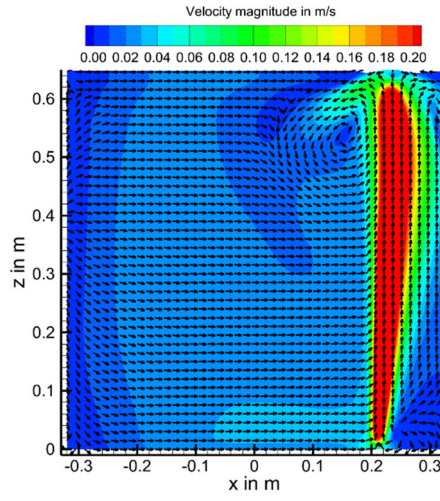
G. Example

The usage of the scoring system is demonstrated with two examples. The first numerical model employs the LES turbulence model with the Germano subgrid model, while the RANS *k*– ϵ model is employed in the second. Apart from that, the same submodels and parameters listed in

Table VIII are used. The subsystem scorings for both models, computed with the programs provided in the database, are summarized in Table IX.

A comparison of the total score shows that the LES, time-averaged over 120 seconds, is more suitable to model the isothermal flow in the ladle. The subsystem scores allow an explanation and reveal potentials for further improvements. As the contour plots suggest, the LES model computes a more reasonable representation of the overall flow structure on the symmetry plane and, consequently, yields a higher toroid score. Both models give a reasonable representation of the mean flow components, but the monitoring points score indicates the potential for further improvement on a more detailed level. Most notable is the low score in the bubble subsystem, indicating a great improvement potential. The main reason for the large deviation is that the real plume oscillates radially, resulting in a flatter but broader profile than the numerical plume.

Table IX. Comparison of Numerical Models by Validation Score

	LES	$k-\epsilon$
<div style="display: flex; justify-content: space-around;"> <div style="text-align: center;"> <p>LES</p>  </div> <div style="text-align: center;"> <p>$k-\epsilon$</p>  </div> </div>		
Toroid		
Total	9.52/10	0.80/10
Mean Velocity		
u	2.58/5	1.93/5
u'	4.37/5	3.20/5
w	5.00/5	5.00/5
w'	4.73/5	3.79/5
Total	16.68/20	13.91/20
Monitoring Points		
$\sum u$	4.23/8	2.37/8
$\sum u'$	6.43/8	2.40/8
$\sum w$	5.80/8	1.49/8
$\sum w'$	5.36/8	3.72/8
$\sum \theta$	7.22/8	3.26/8
Total	29.03/40	13.24/40
Plume Velocity		
\sum_{\max}	5.90/15	3.30/7.5
\sum_{width}	3.17/15	2.47/7.5
Total	9.06/15	5.77/15
Bubble Swarm		
\sum_{mean}	0.68/15	2.71/7.5
\sum_{width}	0.03/15	0.00/7.5
Total	0.71/15	2.71/15
Total	65.01/100	36.42/100

H. Conclusions

For the sake of a strong sense validation benchmark for the isothermal ladle flow, validation experiments were conducted. Validation experiments differ fundamentally from physical modeling experiments because their emphasis is on providing comparable data and quantifying the measurement uncertainties rather than gaining new information about the flow. The data are gathered in a publicly available database that includes additional documentation of the experimental setup and all boundary conditions. By that, the experiment can be replicated numerically with a minimum of assumptions and boundary uncertainties so that a comparable accuracy assessment can be made. For the given case

of an isothermal flow, a validation score is proposed that quantifies the numerical model's accuracy in a single number. Therefore, the models can be compared objectively with measurable quantities, which allows a systematic investigation of influencing factors and an overall optimization of the numerical model. Furthermore, it can serve as a CFD benchmark such as the studies by Odenthal *et al.*^[72,73]

The isothermal flow is only one aspect of the complex ladle system. However, it is a good starting point for a systematic improvement of the numerical model because an accurate modeling of the flow is important for most other subsystems in the ladle like inclusion removal, mixing, or slag entrapment.

To transform the validation database into a strong-sense database in accordance with Oberkampff *et al.*,^[49] some developments are necessary. First is to provide validation data for the missing subsystems such as the upscaling to industrial scales, slag entrapment, thermal quantities, reactions, or inclusion removal. A major challenge will be to develop validation experiments that allow both reproducible, accurate results and an uncertainty analysis. For most, phenomena approaches were published; however, emphasis was on deterministically deriving information about the underlying physics rather than on detailed documentation. An in-depth analysis of whether these experiments are suitable for validation has not been made yet. The methods used in this work can be a guideline for documentation, uncertainty quantification, and publication of these systems. The second missing factor is the general dissemination and acceptance of the database and its use. This requires that guidelines and data are constantly adapted together with users and that complete documentation is maintained.

If these points are supplemented, a strong-sense validation database will be a powerful tool that allows a systematic improvement of numerical models, a wider acceptance of CFD, and finally optimized plant operation practices.

ACKNOWLEDGMENTS

We gratefully acknowledge the support from the German Research Foundation (Deutsche Forschungsgemeinschaft (DFG)) within Project No. RU 2050/2-1. Bubble swarm tracking with BubCNN was performed with computing resources granted by RWTH Aachen University under Project No. nova0032. The authors are grateful to Maria Thumfart (K1-Met, Austria), Hans-Jürgen Odenthal, Norbert Vogl, Fabian Krause, and Andreas Kemminger (SMS Group, Germany), Abhishek Dutta (KU Leuven, Belgium), Mirko Javurek (Johannes Kepler University Linz, Austria), Yu Liu (KTH Stockholm, Sweden), and Ville-Valtteri Visuri and Petri Sulasalmi (University of Oulu, Finland) for their valuable suggestions about the experiments, the reported quantities, and the structure of the database.

OPEN ACCESS

This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence

and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

FUNDING

Open Access funding enabled and organized by Projekt DEAL.

REFERENCES

1. Y. Sahai and R.I.L. Guthrie: *Metall. Trans. B*, 1982, vol. 13B, pp. 203–11.
2. D. Mazumdar and R.I.L. Guthrie: *Metall. Trans. B*, 1985, vol. 16B, pp. 83–90.
3. S. Joo and R.I.L. Guthrie: *Metall. Trans. B*, 1992, vol. 23B, pp. 765–78.
4. Y.Y. Sheng and G.A. Irons: *Metall. Trans. B*, 1992, vol. 23B, pp. 779–88.
5. D. Mazumdar, R.I.L. Guthrie, and Y. Sahai: *Appl. Math. Model.*, 1993, vol. 17, pp. 255–62.
6. D. Mazumdar, H. Nakajima, and R.I.L. Guthrie: *Metall. Mater. Trans. B*, 1988, vol. 19B, pp. 507–11.
7. H. Tse-Chiang, T. Lehner, and B. Kjellberg: *Scand. J. Metall.*, 1980, vol. 9, pp. 105–10.
8. M.B. Goldschmit and A.H. Coppola-Owen: *Ironmak. Steelmak.*, 2001, vol. 28, pp. 337–41.
9. S. Ganguly and S. Chakraborty: *ISIJ Int.*, 2004, vol. 44, pp. 537–46.
10. B. Olika, Y. Pan, B. Björkman, and C.-E. Grip: *Scand. J. Metall.*, 1996, vol. 25, pp. 18–26.
11. J. Aoki, B.G. Thomas, J. Peter, and K.D. Peaslee: *AISTech Proc.*, 2004.
12. C.G. Méndez, N. Nigro, and A. Cardona: *J. Mater. Process. Technol.*, 2005, vol. 160, pp. 296–305.
13. A.H. Castillejos and J.K. Brimacombe: *Metall. Trans. B*, 1987, vol. 18B, pp. 659–71.
14. R. Singh and D. Mazumdar: *ISIJ Int.*, 2008, vol. 48, pp. 1033–35.
15. N.B. Ballal and A. Ghosh: *Metall. Trans. B*, 1981, vol. 12B, pp. 525–34.
16. S.W.P. Cloete, J.J. Eksteen, and S.M. Bradshaw: *Prog. Comput. Fluid Dyn.*, 2009, vol. 9, pp. 345–56.
17. C.A. Llanos, S. Garcia, J. Ángel-Ramos-Banderas, J.D.J. Barreto, and G. Solorio: *ISIJ Int.*, 2010, vol. 50, pp. 396–402.
18. D. Geng, H. Lei, and J. He: *Int. J. Miner. Metall. Mater.*, 2010, vol. 17, pp. 709–14.
19. F.D. Maldonado-Parra, M.A. Ramírez-Argáez, A.N. Conejo, and C. González: *ISIJ Int.*, 2011, vol. 51, pp. 1110–18.
20. C.E. Grip, L. Jonsson, P. Jönsson, and K.O. Jonsson: *ISIJ Int.*, 1999, vol. 39, pp. 715–21.
21. H. Liu, Z. Qi, and X. Mianguang: *Steel Res. Int.*, 2011, vol. 82, pp. 440–58.
22. K. Krishnapisharody and G.A. Irons: *Metall. Mater. Trans. B*, 2006, vol. 37B, pp. 763–72.
23. K. Yonezawa and K. Schwerdtfeger: *Metall. Mater. Trans. B*, 1999, vol. 30B, pp. 411–18.
24. J.W. Han, S.H. Heo, D.H. Kam, B.D. You, J.J. Pak, and H.S. Song: *ISIJ Int.*, 2001, vol. 41, pp. 1165–72.
25. V. De Felice, I.L.A. Daoud, B. Dussoubs, A. Jardy, and J.-P. Bellot: *ISIJ Int.*, 2012, vol. 52, pp. 1273–80.
26. Y. Xie, S. Orsten, and F. Oeters: *ISIJ Int.*, 1992, vol. 32, pp. 66–75.
27. W. Lou and M. Zhu: *Metall. Mater. Trans. B*, 2013, vol. 44B, pp. 1251–63.
28. A.O. Huang, H. Gu, M. Zhang, N. Wang, T. Wang, and Y. Zou: *Metall. Mater. Trans. B*, 2013, vol. 44B, pp. 744–49.
29. Yu. Shan and S. Louhenkilpi: *Metall. Mater. Trans. B*, 2013, vol. 44B, pp. 459–68.

30. J.-P. Bellot, V. De Felice, B. Dussoubs, A. Jardy, and S. Hans: *Metall. Mater. Trans. B*, 2014, vol. 45B, pp. 13–21.
31. L. Zhang and B.G. Thomas: *ISIJ Int.*, 2003, vol. 43, pp. 271–91.
32. L. Li, Z. Liu, M. Cao, and B. Li: *JOM*, 2015, vol. 67, pp. 1459–67.
33. L. Li, B. Li, and Z. Liu: *ISIJ Int.*, 2017, p. ISIJINT-2017-069.
34. T. Haas, A. Rueckert, and H. Pfeifer: *7th Int. Conf. Modelling and Simulation of Metallurgical Processes in Steelmaking (STEELSIM2017)*, Qingdao, China, 2017.
35. H. Duan, L. Zhang, B.G. Thomas, and A.N. Conejo: *Metall. Mater. Trans. B*, 2018, vol. 49B, pp. 2722–43.
36. E.K. Ramasetti, V.-V. Visuri, P. Sulasalmi, R. Mattila, and T. Fabritius: *Steel Res. Int.*, 2018, vol. 90, p. 1800365.
37. E.K. Ramasetti, V.-V. Visuri, P. Sulasalmi, T. Fabritius, T. Saatio, M. Li, and L. Shao: *Metals*, 2019, vol. 9, p. 829.
38. N. Alia, V. John, and S. Ollila: *Appl. Math. Model.*, 2019, vol. 67, pp. 549–56.
39. T. Palovaara, V.-V. Visuri, and T. Fabritius: *7th Int. Congr. Science and Technology of Steelmaking*, 2018.
40. S. Mimouni, S. Fleau, and S. Vincent: *Nucl. Eng. Des.*, 2017, vol. 321, pp. 118–31.
41. R. Denèfle, S. Mimouni, J.-P. Caltagirone, and S. Vincent: *Comput. Fluids*, 2015, vol. 113, pp. 106–11.
42. S.L. Pigny: *Nucl. Eng. Des.*, 2011, vol. 241, pp. 874–87.
43. J. Mietz and F. Oeters: *Steel Res.*, 1989, vol. 60, pp. 387–94.
44. Y. Xie and F. Oeters: *Steel Res.*, 1992, vol. 63, pp. 93–104.
45. Y. Xie and F. Oeters: *Steel Res.*, 1994, vol. 65, pp. 315–19.
46. W.L. Oberkampf and D.P. Aeschliman: *AIAA J.*, 1992, vol. 30, pp. 2000–09.
47. W.L. Oberkampf and T.G. Trucano: *Nucl. Eng. Des.*, 2008, vol. 238, pp. 716–43.
48. W.L. Oberkampf and B. Smith: *52nd Aerospace Sciences Meeting*, 2014, p. 0205.
49. W.L. Oberkampf, T.G. Trucano, and C. Hirsch: *Appl. Mech. Rev.*, 2004, vol. 57, pp. 345–84.
50. T. Haas, M. Eickhoff, and H. Pfeifer: *8th Int. Conf. on Modeling and Simulation of Metallurgical Processes in Steelmaking (STEELSIM 2019)*, Toronto, 2019.
51. J. Yenus, G. Brooks, M. Dunn, Z. Li, and T. Goodwin: *Metall. Mater. Trans. B*, 2018, vol. 49B, pp. 423–33.
52. M. Warzecha, J. Jowsa, P. Warzecha, and H. Pfeifer: *Steel Res. Int.*, 2008, vol. 79, pp. 852–60.
53. T. Haas, A.L. Suarez, M. Eickhoff, and H. Pfeifer: *Towards a strong-sense validation benchmark database for numerical ladle flow models*, Database, Aachen, 2020. <http://doi.org/10.18154/RWTH-2020-03939>.
54. K. Krishnapisharody and G.A. Irons: *Metall. Mater. Trans. B*, 2013, vol. 44B, pp. 1486–98.
55. C. Brückner: *Lecture Series*, von Karman Institute for Fluid Dynamics, Sint-Genesius-Rode, 2000, vol. 1.
56. R. Clift, J.R. Grace, and M.E. Weber: *Bubbles, Drops, and Particles*, Courier Corporation, North Chelmsford, MA, 2005.
57. F. Yucheng and Y. Liu: *Int. J. Multiphase Flow*, 2016, vol. 84, pp. 217–28.
58. D. Bröder and M. Sommerfeld: *Meas. Sci. Technol.*, 2007, vol. 18, p. 2513.
59. A. Ferreira, G. Pereira, J.A. Teixeira, and F. Rocha: *Chem. Eng. J.*, 2012, vol. 180, pp. 216–28.
60. A. Karn, C. Ellis, R. Arndt, and J. Hong: *Chem. Eng. Sci.*, 2015, vol. 122, pp. 240–49.
61. Y.M. Lau, K. Thiruvalluvan-Sujatha, M. Gaeini, N.G. Deen, and J.A.M. Kuipers: *Chem. Eng. Sci.*, 2013, vol. 98, pp. 203–11.
62. D.K. Prasad and M.K.H. Leung: *Digit. Image Process.*, 2012, pp. 71–90.
63. T. Haas, C. Schubert, M. Eickhoff, and H. Pfeifer: *Chem. Eng. Sci.*, 2020, vol. 216, p. 115467.
64. S. Ren, K. He, R. Girshick, and J. Sun: *Advances in Neural Information Processing Systems*, 2015, pp. 91–99.
65. F. Yucheng and Y. Liu: *Meas. Sci. Technol.*, 2018, vol. 29, p. 075206.
66. A. Tomiyama, I. Kataoka, I. Zun, and T. Sakaguchi: *JSME Int. J. B*, 1998, vol. 41, pp. 472–79.
67. N.G. Deen, B.H. Hjertager, and T. Solberg: *10th Int. Symp. Applications of Laser Techniques to Fluid Mechanics*, 2000.
68. K.B. Owusu, T. Haas, P. Gajjar, M. Eickhoff, P. Kowitwarangkul, and H. Pfeifer: *Steel Res. Int.*, 2019, vol. 90, p. 1800346.
69. W. Thielicke and E. Stamhuis: *J. Open Res. Softw.*, 2014, vol. 2, p. 1202.
70. N.G. Deen, J. Westerweel, and E. Delnoij: *Chem. Eng. Technol.*, 2002, vol. 25, pp. 97–101.
71. Niels Gerbrand Deen, J. Westerweel, and Bjørn Helge Hjertager: *5th Int. Conf. on Gas-Liquid and Gas-Liquid-Solid Reactor Engine*, Department of Mechanical Engineering, Aalborg University, Melbourne, 2001.
72. H.-J. Odenthal, M. Javurek, and M. Kirschen: *Steel Res. Int.*, 2009, vol. 80, pp. 264–74.
73. H.-J. Odenthal, M. Javurek, M. Kirschen, and N. Vogl: *Steel Res. Int.*, 2010, vol. 81, pp. 529–41.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.